

```

1 CCTGGAAGTG CCAGGGAGCA CTGGAGGCCA CCCAGTCATG GGGGACACCT
51 TCATCCGTCA CATCGCCCTG CTGGGCTTTG AGAAGCGCTT CGTACCCAGC
101 CAGCACTATG TGTACATGTT CCTGGTGAAA TGGCAGGACC TGTCCGAGAA
151 GGTGGTCTAC CGGCGCTTCA CCGAGATCTA CGAGTTCCAT AAAACCTTAA
201 AAGAAATGTT CCTATTGAG GCAGGGGCGA TCAATCCAGA GAACAGGATC
251 ATCCCCCACC TCCCAGCTCC CAAGTGGTTT GACGGGCAGC GGGCCGCCGA
301 GAACCGCCAG GGCACACTTA CCGAGTACTG CAGCACGCTC ATGAGCCTGC
351 CCACCAAGAT CTCCCGCTGT CCCCACCTCC TCGACTTCTT CAAGGTGCGC
401 CCTGATGACC TCAAGCTCCC CACGGACAAC CAGACAAAAA AGCCAGAGAC
451 ATACTTGATG CCCAAAGATG GCAAGAGTAC CGCGACAGAC ATCACC GGCC
501 CCATCATCCT GCAGACGTAC CGCGCCATCG CCAACTACGA GAAGACCTCG
551 GGCTCCGAGA TGGCTCTGTC CACGGGGGAC GTGGTGGAGG TCGTAGAGAA
601 GAGCGAGAGC GGTGGTGGT TCTGTCAGAT GAAAGCAAAG CGAGGCTGGA
651 TCCCAGCGTC CTTCCTCGAG CCCCTGGACA GTCTTGACGA GACGGAAGAC
701 CCTGAGCCCA ACTATGCAGG TGAGCCATAC GTCGCCATCA AGGCCTACAC
751 TGCTGTGGAG GGGGACGAGG TGTCCCTGCT CGAGGGTGAA GCTGTTGAGG
801 TCATTACAA GCTCCTGGAC GGCTGAAAAG ACGACGTAC AGGCTACTTC
851 CCGTCCATGT ACCTGCAAAA GTCAGGGCAA GACGTGTCCC AGGCCCAACG
901 CCAGATCAAG CGGGGGGCGC CGCCCCGAG GTCGTCCATC CGCAACGCGC
951 ACAGCATCCA CCAGCGGTCT CGGAAGCGCC TCAGCCAGGA CGCCTATCGC
1001 CGCAACAGCG TCCGTTTTCT GCAGCAGCGA CGCCGCCAGG CGCGGCCGGG
1051 ACCGCAGAGC CCCGGGAGCC CGCTCGAGGA GGAGCGGCAG ACGCAGCGCT
1101 CTAAACCGCA CGCGGCGGTG CCCCCGCGC CGAGCGCCGA CCTCATCTG
1151 AACCCTGCA GCGAGAGCAC CAAGCGGAAG CTGGCGTCTG CCGTCTGAGG
1201 CTGGAGCGCA GTCCCCAGCT AGCGTCTCGG CCCTTGCCGC CCCGTGCCTG
1251 TATATACGTG TTCTATAGAG CCTGGCGTCT GGACGCCGAG GGCAGCCCCG
1301 ACCCTGTGCC AGCGCGGCTC CCGCCACCCT CAATAAATGT TGCTTGAGT
1351 GGAAAAA AAAA AAAA AA
(SEQ ID NO: 1)

```

#### FEATURES:

5'UTR: 1 - 37  
 Start Codon: 38  
 Stop Codon: 1196  
 3'UTR: 1199

#### Homologous proteins:

##### Top 10 BLAST Hits:

Sequences producing significant alignments:	Score (bits)	E value
CRA 18000004925255 /altid=gi 4557785 /def=ref NP_000256.1  neut...	789	0.0
CRA 18000005124568 /altid=gi 2754713 /def=gb AAB95193.1  (U5783...	788	0.0
CRA 18000005207006 /altid=gi 4263750 /def=gb AAD15422.1  (AC004...	783	0.0
CRA 18000005171728 /altid=gi 6685673 /def=sp O77774 NCF1_BOVIN ...	684	0.0
CRA 148000004473069 /altid=gi 8439513 /def=dbj BAA96544.1  (ABO...	670	0.0
CRA 118000005118410 /altid=gi 9623382 /def=gb AAF90134.1 AF2677...	663	0.0
CRA 18000005141875 /altid=gi 3061282 /def=dbj BAA25649.1  (AB00...	659	0.0
CRA 18000005020732 /altid=gi 1171669 /def=sp Q09014 NCF1_MOUSE ...	655	0.0
CRA 18000004937799 /altid=gi 2118398 /def=pir I54525 leukemia-...	651	0.0
CRA 148000001425618 /altid=gi 7839599 /def=gb AAF70344.1  (AF26...	527	e-148

#### EST:

Sequences producing significant alignments:	Score (bits)	E value
gi 12896059 /dataset=dbest /taxon=960...	1532	0.0
gi 12951967 /dataset=dbest /taxon=960...	1501	0.0
gi 12342004 /dataset=dbest /taxon=96...	1423	0.0

#### EXPRESSION INFORMATION FOR MODULATORY USE:

gi|12896059 placenta  
 gi|12951967 B cells from Burkitt lymphoma  
 gi|12342004 primary B-cells from tonsils

#### Tissue Expression:

Leukocyte

FIGURE 1

```

1  MGDTFIRHIA LLGFEKRFVP SQHYVYMLV KWQDLSEKW YRRFTEIYEF
51 HKTLEKMFPI EAGAINPENR IIPHLPAKW FDGQRAAENR QGTLTEYCST
101 LMSLPTKISR CPHLLDFKV RPDDLKLPD NQTKKPETYL MPKDGKSTAT
151 DITGPILQT YRAIANYEKT SGSEMALSTG DWVEVVEKSE SGWWFCQKA
201 KRGWIPASFL EPLDSPDETE DPEPNYAGEP YVAIKAYTAV EGDEVSLLEG
251 EAVEVIHKLL DGWKDDVTGY FPSMYLQKSG QDVSAQRQI KRGAPPRSS
301 IRNAHSIHQR SRKRLSQDAY RRNSVRFLQQ RRRQARPGPQ SPGSPLLEEER
351 QTQRSKPQPA VPPRPSADLI LNRCSESTKR KLASAV
(SEQ ID NO: 2)

```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
N-glycosylation site

131-134 NQTK

[2] PDOC00004 PS00004 CAMP\_PHOSPHO\_SITE  
CAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 4

```

1      42-45 RRFT
2      297-300 RRSS
3      313-316 KRLS
4      321-324 RRNS

```

[3] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

Number of matches: 10

```

1      36-38 SEK
2      53-55 TLK
3      133-135 TTK
4      160-162 TYR
5      300-302 SIR
6      311-313 SRK
7      324-326 SVR
8      352-354 TQR
9      377-379 STK
10     378-380 TKR

```

[4] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 11

```

1      53-56 TLKE
2      93-96 TLTE
3      148-151 TATD
4      171-174 SGSE
5      178-181 STGD
6      208-211 SFLE
7      215-218 SPDE
8      238-241 TAVE
9      246-249 SLLE
10     279-282 SGQD
11     344-347 SPLE

```

FIGURE 2A

[5] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 3

1 83-88 GQRAAE  
2 172-177 GSEMAL  
3 280-285 GQDVSQ

Membrane spanning structure and domains:  
NO DATA

BLAST Alignment to Top Hit:

>CRA|18000004925255 /altid=gi|4557785 /def=ref|NP\_000256.1| neutrophil  
cytosolic factor 1; Neutrophil cytosolic factor-1 (47kd);  
p47phox [Homo sapiens] /org=Homo sapiens /taxon=9606  
/dataset=nraa /length=390  
Length = 390

Score = 789 bits (2015), Expect = 0.0

Identities = 385/390 (98%), Positives = 386/390 (98%), Gaps = 4/390 (1%)  
Frame = +2

Query: 38 MGDTFIRHIALLGFEKRFVPSQHYVYMLVKWQDLSEKWYRRFTEIYEFHKTLKEMFPI 217  
MGDTFIRHIALLGFEKRFVPSQHYVYMLVKWQDLSEKWYRRFTEIYEFHKTLKEMFPI  
Sbjct: 1 MGDTFIRHIALLGFEKRFVPSQHYVYMLVKWQDLSEKWYRRFTEIYEFHKTLKEMFPI 60

Query: 218 EAGAINPENRIIPHLPAKWFDDGQRAAENRQGLTEYCSTLMSLPTKISRCPHLLDFFKV 397  
EAGAINPENRIIPHLPAKWFDDGQRAAENRQGLTEYCSTLMSLPTKISRCPHLLDFFKV  
Sbjct: 61 EAGAINPENRIIPHLPAKWFDDGQRAAENRQGLTEYCSTLMSLPTKISRCPHLLDFFKV 120

Query: 398 RPDDLKLPDNTQTKKPETYLMPKDGKSTATDITGPILQTYRAIANYEKTSGSEMALSTG 577  
RPDDLKLPDNTQTKKPETYLMPKDGKSTATDITGPILQTYRAIA+YEKTSGSEMALSTG  
Sbjct: 121 RPDDLKLPDNTQTKKPETYLMPKDGKSTATDITGPILQTYRAIADYEKTSGSEMALSTG 180

Query: 578 DWVEWEKSESQWFCQMKAKRGWIPASFLEPLDSPDETEDPEPNYAGEPYVAIKAYTAV 757  
DWVEWEKSESQWFCQMKAKRGWIPASFLEPLDSPDETEDPEPNYAGEPYVAIKAYTAV  
Sbjct: 181 DWVEWEKSESQWFCQMKAKRGWIPASFLEPLDSPDETEDPEPNYAGEPYVAIKAYTAV 240

Query: 758 EGDEVSLLEGEAVEVIHKLDDGW----KDDVTGYFPSMYLQKSGQDVSQAQRQIKRGAPP 925  
EGDEVSLLEGEAVEVIHKLDDGW KDDVTGYFPSMYLQKSGQDVSQAQRQIKRGAPP  
Sbjct: 241 EGDEVSLLEGEAVEVIHKLDDGWWIRKDDVTGYFPSMYLQKSGQDVSQAQRQIKRGAPP 300

Query: 926 RRSSIRNAHSIHQSRKRLSQDAYRRNSVRFLQRRRQARPGQSPGSPLEERQTQRSK 1105  
RRSSIRNAHSIHQSRKRLSQDAYRRNSVRFLQRRRQARPGQSPGSPLEERQTQRSK  
Sbjct: 301 RRSSIRNAHSIHQSRKRLSQDAYRRNSVRFLQRRRQARPGQSPGSPLEERQTQRSK 360

Query: 1106 PQPAVPPRPSADLILNRCSESTKRKLASAV 1195  
PQPAVPPRPSADLILNRCSESTKRKLASAV  
Sbjct: 361 PQPAVPPRPSADLILNRCSESTKRKLASAV 390 (SEQ ID NO: 4)

FIGURE 2B

Hmmer search results (Pfam):  
HMM results:

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
CE00053	CE00053 max_mitogenic_oxidase	573.6	1.3e-168	1
PF00787	PX domain	119.4	6.6e-32	1
PF00018	SH3 domain	107.5	2.7e-28	2
CE00036	CE00036 androstane_receptor	0.3	4.6	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00787	1/1	4	121 ..	1	147 []	119.4	6.6e-32
PF00018	1/2	159	213 ..	1	57 []	66.7	1e-16
CE00036	1/1	243	257 ..	199	213 ..	0.3	4.6
PF00018	2/2	229	279 ..	1	57 []	41.0	1.2e-09
CE00053	1/1	1	386 []	1	566 []	573.6	1.3e-168

FIGURE 2C

1	TACTAAAAAT	ACAAAATTAG	CCAGGCGTGG	TGGCGCACAC	CTGTAATCCC
51	AGCTACTTGG	GAAGCTGAGG	CAGGAGAATC	GCTTGAACCT	GGAAGGCAGA
101	GGTTGCAGTG	AGCCGAGATT	GTGCCACTGC	ACTCCAGCCT	GGGCAACAAG
151	AGCGAAACTT	CGCTTCAAAC	AAATAAATTA	ACGCCACAGCA	TGTCTTTGGCT
201	TTCACTGCCC	AGACCTCAAC	CCTCACCCCC	AGGAGATCAG	GTCCGGACCA
251	TGAGCTGACC	CTGGACTCAG	GCAAGGGTGA	GTTGGTGACG	CCCTGGCCTG
301	CTGGGAGGCA	CAGGCTGCAG	CAGGCTGCCT	GGGGCTGAGG	CCCGCCACTC
351	ATGAACTCAT	GACCTTGAAT	GAGCTCCAAA	AGCTCTGGGC	CTCCCAGGCT
401	CTAGGGGGAG	TGGGAGAGAG	AGGCCTCAGC	CTGTCCCTGG	GCATGCTGCC
451	CCCTCCTCAC	CTCTTTGTCC	CAAATCCCCT	TCCTGGCAAA	GCTGACAGTC
501	TTAATATCAC	TCTGGAGAAA	ACTGAGTCAG	CCCTAAGGAA	CAATTCAATG
551	AACCATTTGC	TTACTTGAGG	ATTGGAATC	AAGTCTCACT	CAAAGTCTGT
601	GCCATTTTCG	TCCCAGCTGT	CACTGGCCCT	CATCCACACA	CACCCAAGGA
651	TGAGCATCTA	ACGCTTGCAT	GCACACTCCC	ATGCCCGCGT	TCATTCACTC
701	ATTCATTTCAT	TCATTCACTC	ATTCATTGAC	TCATTCAATC	ATTCACTCAC
751	TCATTCAATC	ACTCAGTGAA	TGTTGCAGTC	ACGATCCAAA	TATTTATGGC
801	CTCTGTGTGC	CAGGCACTAG	ATGGAGGGGC	TGGGGCTAGA	GCCCCTGATA
851	ACCCGGTCAT	GCCCTAGCTT	TCCTGGGACA	CACATTGTGG	TAAGGGGAGA
901	CTAAAAAAT	TAAGTCAGGC	CAGGCACGGT	GGCTCATGCC	TGAATCCCAG
951	CACTTTGGGA	GGCCGAGGCG	AGTGAATTAC	CTGAGGTCAG	GAGTTCAAGA
1001	CCAGCCTGGC	CAACATGGAG	AAACCCAGTC	TCTAATTAAA	AAAAAAAAAA
1051	AAATTAGCCA	GGTGTGGTGG	CACATGCCTG	TAATCCCAGC	TACTCAGGAG
1101	ACTAACGCAA	GAGAATTGCT	TGAACCCAGG	AGGCAGAGGT	TGCGGTGAGC
1151	CGAGATCGCG	CCATTGCACT	CCAGCCTGGG	AAACAAGAGC	GAGACTCCAT
1201	CTCAAAAAAA	AAAAAAGTGG	GAGGCAGAGG	CAGGAGGATC	ACTAGAGGCC
1251	AGTAGTTTGA	GACCATCCTG	GGCAACATAG	CAGGACCCCTG	TCTGTACAAA
1301	AAAATTAAAA	AAAATTTAAC	CGGGCATGGT	GGCACACACC	CGTAGTCCCA
1351	GCTACTCCAG	AGGCTGAGGC	AGGAGGATCG	CTGGAGCCCA	GGAGTTGGAG
1401	GCTGCAGTGA	ACTGTGATCC	CACCACTGCA	CTTAAGCCTG	GATAACAAAG
1451	CAAGACCCTG	TCTCAAATAA	CAATAGCAAT	AATAATAAAG	AAAAATTAAA
1501	TGCAATTTGC	GATGCATCAG	TGATAAGTGC	TCTGCAGAAA	AAGGAGGCAG
1551	GAAGAGGCTG	AGAAAGGTAT	GAGGTTTGCT	ATGCAATGTG	AAGTTATCAA
1601	GGAAGGCTTC	TGGGAAGAGG	TGACATTTGA	GCAGAGAAAT	GGAGGAGAGT
1651	TATGGAGGGA	AGATGGTGAA	TGGGGGGAAC	ATGGTCAAGA	CCAGGAATAT
1701	GGTCAAGGGG	GGAAAGATGG	TCAAGGGGAC	GCAGCAAATG	CAAAGGCCCT
1751	GAGGCAGGAG	CAGCTTGATT	CACCCCCAAA	ACCCGTGGGG	CCCGTGCAGG
1801	CGACGGGAAG	GACAAGTGTA	AACCCCTTTC	CTTGTCCCTG	CAGGTGTGTG
1851	TGAACATGAG	TCTGCCCATG	TTTACACCCT	GCAAGCCTGA	AGAGTCCCCA
1901	GAAACTGAAA	GAGAAGCAA	AGCCCTTTCT	GTACCCTCCC	TGCCCCCTGT
1951	CCCGACCGCG	ACAAAAGCGA	CTTCCTCTTT	CCAGTGCAAT	TAAGGCGCAG
2001	CCTGGAAGTG	CCAGGGAGCA	CTGGAGGCCA	CCAGTCATG	GGGGACACCT
2051	TCATCCGTCA	CATCGCCCTG	CTGGGCTTTG	AGAAGCGCTT	CGTACCCAGC
2101	CAGCACTATG	TGAGTAGCTG	GTGGAGGGCA	TCCCGTGGGG	GGAATACGGG
2151	AGGGACAGCA	CGGCCACCCT	TGCAGTCCCA	GGGCCAACCA	GCTCCAGTGA
2201	GGACTAACGG	GGCAGGGTCT	TGGGCACCTG	GTCCCTGGTC	TTTGAGCCTG
2251	GATCTACCCC	TCTGATCCCT	GGGAAGACAG	TTCCCTTGGA	CCCGCCCTGG
2301	GCCCCAGCCC	TTTACTGTCC	CCGCCTGTGT	CCCCAGCCAG	GCCCTCAGCC
2351	TTAGCCAGGA	GTCCTCTTTC	TGCTCCCCTG	CCATGGCCAG	GCAGCCCAGC
2401	GCTCTCTCAG	GTCCGAGGCC	CACTCCTCCA	GGAAGCCTTC	CCTGACTAGC
2451	CCAGCTATCA	GAGAGTGGCC	CTCCCAAGAG	GGAGGCCTGG	AAACTAAAGC
2501	TCTCTCTCTC	CCCAGCTGCC	TGTAGTGTCA	GTTAGAGTCT	TATCCTCTCC
2551	AGTAGGGTGA	CACCATGACA	GGGGCCAATA	GAGTCCTCCC	ATCTGTCCCC
2601	AAGGAGGCTG	GACAAATGCC	TGCTCAGACA	CACAAGTCCA	CTGGGTCCCC
2651	TAATCCCAT	GGAAGGCCAG	GGAGGAACTA	CATTTAGGAA	ATTGAAGCTT
2701	GTATGGAACA	TTTAGTCCTA	TGTGCCAAGA	CCTTTCTCTT	TTTTGTATT
2751	TTTTTGTGTT	TTGAGACAGA	GTCTTGATCT	GTTGCCCAGG	CCAGAGTGCA
2801	GTGGCAGCAT	CTCAGCTCAC	TGCAACCTCC	GCCTTCCAGG	TTCAACTGGT
2851	TCTCCTGCCT	CAGCCTCCAG	AGTAGTTGGG	ATTACAGGTG	CCCACCACCA
2901	CGCCTGGCTA	ATTTTTGTAT	TTTTAGTAGA	GACAGGGTTT	CACCATGTTG
2951	GCCAGACTGG	TCTCAAATC	CTGACCTCAA	GTGATCCACC	CACCTGGGCC
3001	TCCCAAAGTG	CTGGGATTAC	AGGCATGAGC	CACCGTGCCT	GGCCTGTTTT
3051	TTTGAAATGA	GGTCTGGAGT	GCAGTGGTGC	GATCATAGTT	CACTGCAGCC
3101	TCAACCTCCC	AGGCCCAAGT	GATCCTCCTG	CCTCAGCCCC	TTGAGTAGCT
3151	GGGGCTACAG	GCGCACACCA	CCATGCCTGG	CTAGTTTTTA	AAATTTTTGT
3201	GGAGATGAGG	TTTCACTATG	TTGTCCAGGC	TAATCTTGAA	CTCCTCGGCT

FIGURE 3A

3251	TAAGCAACCC	TCTGGTCTCA	GCCTCCCACA	GTGCTAGGAT	TACAAGCGTG
3301	AGCTACCGTG	CCTAGTCACT	TTTCTCCTTT	TCTTTGTAA	TTTCAGTTTT
3351	GAAATTTCAA	ATTTACAGAA	AGGCTACTGG	GTGTCAAAAC	GGTACCAGTC
3401	ACTCCAATAG	TCCTTCACTC	ACCTTCATCC	ACACCTCTCT	TTCTGGGGAT
3451	ATTTTCTGAA	TTATTTGAGA	GTGAGTTGAA	GACGTGTTTC	TTTACCTCTA
3501	AATACTAGTT	GTTGGGCATT	TCTTAAAATC	AAGGCATTCT	CTTACATAAT
3551	CACAACACAC	GTGTCAAAAT	CAGGAAATTA	ACATGGACAA	AACACCATTA
3601	TCCACCCACA	GACTTTACTG	AGGTTTCCCC	GATTATCCTG	CTTGTCTCT
3651	GCAGTGAAAA	CTTTTTTCAG	GTCTAGGATC	CAGTCAAGGA	TCAATGTCAT
3701	AGCCTTTAAC	CTTCTTTAAT	CTGGATCAGT	CTTTTTTCTT	TTTCTTTTTC
3751	TTTTTTTGA	CACGGAATCT	CACTCTGTCT	CCAGACTGGA	GTGCAGTGGT
3801	GCAATCTCGG	CTCATTGCAA	CCTCTGCCTC	CTGGGTTCAA	GAGATTCTCC
3851	TGCCTCAGCC	TCCTGAGTAG	CTGGGAATAC	AGGTCCGCGC	CACCACGCCC
3901	AGCTCGTTTT	TGGTAGAGAC	AGGTTTTTGC	CATTGATTCT	GGATCAGTCT
3951	TTTTTTTTTT	TTTTATGAGA	TGGAGTCTTA	CTCTGTCACC	CAGGCTGGAG
4001	TGCAATGGCA	CAATCTCCAC	TCACTGCATC	CTCCGCCTCC	CAGGTTCAAG
4051	CAATTCTCGT	GCCTCAGCCT	CCCGAGTAGC	TGGGATTACA	GGCATGCGCC
4101	ACCATGCCCG	GCTACTTTTT	GTATTTTTAG	TAGAGACAGG	GTTTCACCAT
4151	GTTAGCCAGG	CTGATCTCGA	ACTCCTGACG	TCAGGTGATC	TGCCCCCCTC
4201	GACCTCCCAA	AGTGCTGGGA	TTACAGGCGT	GAGCCACCGT	GCCAGCGGAT
4251	TCTGGATCGG	TCTTAATCAG	TCTTTGTCTT	TTGCAACTTT	GATGTTTTGC
4301	AGAGAGCAGA	CCAGTTACCT	TGTAGAATGT	CCCTTAGATT	GGGTTTATCT
4351	TCATTAGATT	CAGTTTGTGT	ATCCAGGGCA	GTGGATCTTA	GATGCAATTC
4401	TGCTTCTTTT	TTAATTTTTT	TGAGAGGGAG	TCTCGCTCTG	TCACCCAGGC
4451	TGGAGTGACG	TGGCACAAAC	TCAGCTCACT	GCAGCTCCG	CCTCCCGGGT
4501	TCAAGCAATT	CTCCTGTCCC	AGCCTCCCAA	GTAGCTGGGA	TCACAGGTGC
4551	CCATCACCAC	TACCGGGTAA	TTTTTGTTT	TTTAGTAGAG	ACAGGGTTTC
4601	ACCATATTGG	TCAGGCTGGT	CTTGAACGCC	TGACCTCAGG	TGATCCACCT
4651	GCCTTGGCCT	CCCAAAGTGC	TGGGATTACA	GACGGGAGCC	AACATGCCCA
4701	GCCTTCTGTC	CCCTCCCGTC	CCCTCCCCTC	TCTCCTGTC	CCCTCCCTTC
4751	CCCTCCCTTA	TCCTCATGTC	CCCTCCCCTC	CCCTCCCCTC	CCCACCCAAG
4801	CTGGAGTGCA	GTGGTGCAAT	CATAGCTCAC	TAAAGCCTTG	ACCTCCAAGT
4851	CTCAAGCAAT	TCTCTGCCT	CACCTGGGGC	CACAGGTGTG	CGGCACCACA
4901	CCCGGACAAT	TTTTGTGTTT	TTAGTAGATA	TGGGGGTCTC	GCTATGTTGC
4951	CCAGGCTGGT	CTCAAACCTCT	TGGACTCAAG	CGATCTTCCC	ACCTCGGTAC
5001	TAAAAAGTGC	TGGGATTCCA	GGTGTGAGCC	ACCGTGCCCA	GCCTAGGTCC
5051	TACTTTTATC	TCCAATTTAC	AGATGAGTCC	ATTTGAGAGA	AGCTGACCCT
5101	CTTGCCCTGG	GTCTCAAGGC	TGGGGCGTGG	CAGCACTTGG	GTCCACGTTT
5151	GTGCCCTTTC	TGCAATCCAG	GACAACTGCA	AAGATGGTCC	TCACCCCAAT
5201	CCTCTGGGCT	TCCTCCAGTG	GGTAGTGGGA	TCCTGGGTGC	ACACAGCAAA
5251	GCCTCTTTGG	AGGCTGAATG	GGGTCCCCCG	ACTCTGGCTT	TCCCCCAGGT
5301	ACATGTTTCT	GGTGAAATGG	CAGGACCTGT	CGGAGAAGGT	GGTCTACCGG
5351	CGCTTCACCG	AGATCTACGA	GTTCCATGTG	AGTGTGGGGA	CGGAGGAGGG
5401	ACAGGGACCC	ACCGTTCCAG	CTCCACCCTT	TGGGAAGGAC	CTTAGCCAG
5451	GTGATGGGGA	AACTGCAGAA	CCCAGAATCC	CCTCCCAGAC	CACAGTTAAA
5501	GGGGATTTAT	TTAGTTATAT	AAATTTTTGT	GACAGGGTCT	TGCTCTGTCA
5551	CCCAGGGTCT	TGCTCTGTCA	CCACTCTGAA	CACCTCATGT	TCTCTGATTA
5601	CAGGCATGAG	CCCCACGGT	CGGCTTTTAA	GGTGGTTTTG	AGAGGTATTT
5651	AGGTTTGACG	TGCAGGGGCG	CAATCATAGC	TCACTGCAGC	CTCAACCTCT
5701	GGGGCTCAAG	CGATCCTCCT	GCCTCAGCCT	CCTGAGTAGC	TGGGACTATA
5751	GGTGCGCATC	ACCATGTGTG	GCTAATTTTT	GTATTTTTTA	TAAAGATGGG
5801	GATCTCACTA	TGTTGCCAG	GCTGGTCTTG	AACTCCAGAC	CTCAAGTGAT
5851	CCTCTGCCT	TGGCCTCCCA	AAGCTAAGGG	GGCATTAAAA	GAAAAAAACA
5901	TTTTTCCCCC	TGAAACATTT	AAGTAGTCTT	ACTGAAAAACA	ATAAAACACA
5951	GAAACACCAG	ATTCTCATTT	TAAAGTAAAA	CAGACAGGAT	CTCCCAGAAC
6001	CTTCTAGAA	TGGAACCATT	CTTGTCGCTT	TTGAAAAACA	AAGCCAAGTT
6051	CTAGATCCCC	AATAAATGCA	CCTGCTGGTG	AACATTCTCC	TTGTGGTTCT
6101	CGTCCCTATG	TTAGTTATTT	TCCTAAATTT	TACATTTGTA	CCTTTTTAAG
6151	AATGAGTTAT	CAGTTTTTTT	ATATTTGCTT	TTCTTTTGAG	ATGGGGTCTT
6201	GCTCTGTCAC	CCAGGCTGGG	GTGCAGTGGT	GCAATCACGG	CTCACTGCAG
6251	CCTCAACCTC	CAGGGCTGAA	GCGATTCTCC	CATCTCAGCC	TCCCATGTTG
6301	AGATCACAGG	TGTGCACCAC	CACACCTGGC	TCCTTTTCTT	GATTTGTTTT
6351	TTGTAGAGAT	GGGATTTCCG	TATGTTGCCC	AGGCTGGTCT	CTAACTCCTG
6401	GACTCAAGTG	ATCCTCCCGC	CTCAGCTTCC	CAAATTTGCTA	GGATTACAGG
6451	TTTGAGCCCC	TGCACCTGGT	CAACCTGAGT	TTTAAGAGGA	TCCCTTTGGC

FIGURE 3B

6501	GACTGGATTG	AGGACAGACA	AGAGTGGACG	GGGGACACAA	GGAGGCCATT
6551	TTGCTTATCC	AGGCCTGGTA	GTGGCTAGGG	CCAGGAGGGT	GGGGTTGGTG
6601	GGAAGCAGTC	AGATCCCAAA	GAGATTTGGG	GATTGGAAGC	AAAAGGATTT
6651	GCTGGTGACT	TGCACATGGG	AGGGAGAGAG	GTCAGTGCCT	CTGTTAATCA
6701	AGGAATCCAG	ATTGCCACCG	AAATTTCTAG	GCCCCAGATA	TTTAGGTAGT
6751	GTCTCACTCT	GTACCCAGG	ATGGAGTGCA	GTGGCGCCAT	CTCGGCTCAC
6801	TGTAACCTCC	GCCTCCAGG	TTTAAACGAT	TCTCCACCT	CAGCCTCCTG
6851	AGTAGCTGGG	ATTACAGGCA	TGTGCCACCA	CTCCCGGCTA	ATTTTTGTAT
6901	TTTTAGTAGA	GACGGGGTTT	CACCACGTTG	GCCAGGCTGG	TCTTGAACTC
6951	CTGACCTCAA	GTGATCCACC	CACGACAGCC	TCCCAAAGTG	CTGGGATTAC
7001	AGGCGTGAGC	CACCATGCTC	GGCCTTTTAG	GTGGTTTTGA	GAGGTATTTA
7051	GGTCACTTCC	AATCTCGTGC	TTTTCCAAGT	GTTGTAAACT	ACAAATATTC
7101	CTTCACGTCT	TCTTGTCTTT	TTAATGTTTA	GAAAACCTTA	AAAGAAATGT
7151	TCCCTATTGA	GGCAGGGGCG	ATCAATCCAG	AGAACAGGAT	CATCCCCCAC
7201	CTCCCAGGTG	AGCACGGGGC	TGAGCCGCT	GTACGGGGT	CATTGGCGGG
7251	GCTCAGCTG	CCCTCCAGC	ACCTCTCGGG	CTTGACCTCA	TGTTCTCTGG
7301	TGCCAGCTCC	CAAGTGGTTT	GACGGGCAGC	GGGCCGCCGA	GAACCAACAG
7351	GGCACACTTA	CCGAGTACTG	CAGCACGCTC	ATGAGCCTGC	CCACCAAGAT
7401	CTCCCCTGT	CCCCACCTCC	TTGACTTCTT	CAAGGTGCGC	CCTGATGACC
7451	TCAAGCTCCC	CACGGACAAC	CAGTGAGTGA	ACTTTTCACC	CTGCCAGGTG
7501	GGAGAGGGAA	GGAGGGGTGG	GACTTCTGT	GTTTTGCAGA	TGAGGAAACC
7551	AAGGCTCAGA	GAGGGAAGC	CACCTTCCCA	GAGCCACACA	GCCGAAAGA
7601	GGAGGCAAAAT	TCCACCTCCG	GCCCTGTGA	CCCCGCCAAG	CCTGCACCTT
7651	AATCTTTCAC	ACCTCAGGGC	ACTGGGGGAA	GCACTCGGGG	CTGGAGGTTT
7701	AAAGTCCTGG	GTCTCATCC	TGACATTATG	GCCACCTGGC	CATGGGACCT
7751	GGAGCCAGTC	ACCACTGCTC	TCTGAATGCA	GGTTGTCCAT	TTCTATAATG
7801	GGCAGTGAGG	ATCAGATGAA	GCATTGGGTG	TCTTGCGGAG	CCCCCAGAA
7851	GGATGTGGGG	TTGATGCCCT	TGCTAAGTGC	TGAGCATGTC	TGGGGTCTCC
7901	TGTACCCAGG	ACCCTGTGTG	GAAGGCACCT	GAGAGGCTGA	GGGAGCTCCA
7951	GGCAGGCTGG	GGAAGTCCCC	TTCTCCACTC	CTCTCTGGTC	ACTGAAGCTC
8001	GAAGTGGGGA	GCATGAGGAC	AGGACGTTAC	CCCTTGTCAA	GGCACCCAGG
8051	CTGCCAAGAC	AGAGACAAGC	AGCATTGCTC	CGGCCAGCAC	TTATTGACGC
8101	TTGAAGGTGT	CCCCTGGCCC	AAGGAAGGGC	AGTTATCATC	AGCCCCGGAG
8151	GCGGGGGAAG	GATGGACTCT	GCAGTGGGGT	CCGCTCCTCA	TTGCCCTGCTC
8201	TCTCAGGGCT	CCAGAAGGAG	GAAGAGGCCG	GGCACAGTGG	CTCACACCTA
8251	TAATCCCAGC	ACTTTGGAAG	GTCGAGGTGG	GCAGATCACC	TGAGGTTGGG
8301	AGTTTGAGAC	CAGCCTGGCC	AACATGGTGA	AACCCCATCT	CTACCAAAAA
8351	TATAAAAAAT	TAGTCAGGCA	TGGTGGTGTG	CGCTTGTAAT	CCCAGCTACT
8401	TGGGAGGCCG	AGGCAGGAGA	ATCGCTTGAA	CCCCGGAGGC	AGAGGTTGCA
8451	GTGAGCCAC	ACTGCCAC	TGCACTCCAG	CTGGGTGAC	AGAGCGAGAC
8501	TCTGTCTAAG	AAAAAAAAAA	GAAAAAGA	AAGAAGATGG	CCTGGGAGCC
8551	CGCAAGAGCA	TTTTCCAGGC	TTAGGGCATC	CTTTGGGTCT	GCAGAAGGCT
8601	ATGCAGTGTG	CTCTCATGT	CCCTCCCTTG	GGCTGCCCGA	GCAGATCCGC
8651	CCGCCCCCAT	CACCTCCTGA	AGCCCTTCT	CAGCCAGTCC	AGTTGCTGTC
8701	TTCTCTCCCG	AGTGCCCTTT	CCCTTTCCCG	GGTCCCTCTT	CTCTTGGGAA
8751	GTTCTTCTGC	AGGTCTACCC	AGTGCCTCTT	CTTCTCCTAT	GGGAAGCCAA
8801	GGGTCTCACC	CAGACTGTTT	TCTCCTCAGG	ACAAAAAAGC	CAGAGACATA
8851	CTTGATGCCC	AAAGATGGCA	AGAGTACCGC	GACAGGTGAG	AGGACGGGGG
8901	GCAGCCGCG	GGGGGGGACA	CCCTGAGGAG	ACCCAGAGTG	TTCAGGGAAT
8951	GGAGCAGGGG	CTGGGAGCAG	GCTGGGAGGG	CTCACAGCTA	CCCTGCTGAA
9001	GAATTGGGTG	TTTGGGCCGG	GTGCGGTTGC	TCATGCCTGT	AATCCCAGCA
9051	GTTTGGGAGG	CCGAGGCAGG	TGGATCACTT	GAGGTCAGGA	GTTTGAGACC
9101	AGCCTGGCCA	ACATGGAGAA	ACCCTGTCTC	TACTAAAAAT	CCAAATTAGC
9151	CAGGCGTGGT	GACAGGTGCC	TGTAGTCCCA	GCCACTTGGG	AGGCTGAGGC
9201	AGGAGAATTG	CTTGAACCCG	GAAGACGGAG	TTTGCAGTGA	GCCGAGATCG
9251	TGCCACTGCA	CTCCAGCCTG	GGCAGCAGAG	CCAGACTCCA	TCTCAAAAAA
9301	AAAAAAAAAA	AAGAAGAATT	GGGTCTTTGG	AAGGTCCCTG	GAGACTGAAA
9351	GGAGCCCTTT	GCAGGTGGCA	GTGCAGAGAC	CAGCGCAGAC	CCTTGCTACT
9401	GGCAGCCGGG	GGAGTGTGTT	CGGCTGAATG	AATGAACAGG	TTTTGGAGGG
9451	CAGCGTGGCC	TTGAGAGGCG	ATGCAGGGCT	GTGGCAGTTT	CTAATACTTA
9501	TTGCACAGTC	ACTGCTAATA	ACAATAATAA	TAATAATACC	TAACATTAAT
9551	GGAGTGCTTA	CTCTGTGCCA	GCCACTATTT	TGTTTTTGTT	GTTTTCAGTG
9601	ACAGGGTCTC	GCTCTGTTGC	CCAGGCCAGA	GTGAAGTGGT	GTGATCATAG
9651	CTCACTACAG	CCTCGACCTC	CTGGGCTGAA	GCGATCCTCC	CACCTCAGCC
9701	TCCCAAGTAG	CTGGGATTAC	AGGTGTGTGC	CACCATGTCC	AGCTAATTTT

FIGURE 3C

9751 TAATTTTCTG ATAGAGATGG GGTCTCACTA CATTGCCCAG GCTGGTCTTA  
 9801 AGCTCTTGGC CTCAAGCAAC CCTCCTGCCT CAGCCTCCCA AAGTGCTGAG  
 9851 ATTATAGACA TGAGCCACTG TGCCCGGCTT TTTCTTCTTC TTATAAGGAC  
 9901 ACGAGGCTCG TTGGGTTAGG GCCCACTCTA CTGACCTCAT TTTAACTTAA  
 9951 TTACCTCTTG AAACGTACTT AAGAGTACCT TTCTCTTAAT ACACCCACAC  
 10001 TGTAAGGTAC TGGGTGGTTA GGACTTCAAC ATATGAATTT TGAGAAGGCG  
 10051 GATGTCAGCC AATACCAAAC AGCATCAGCA CCTCCACGGT TGGATGAAGG  
 10101 GCTGTCAGA AATGCACACT CAGGTCCAC AGTGACCTA CTGAACAGGA  
 10151 TAGGCATTTT AGCAAAATCC CAGGTATTCC GGTGCACCTT AAAGTTAGGA  
 10201 AAAGGTCAGG CACTGTGGCT CATGCCGTGA ATCCCAAGC TTTGGGAGGC  
 10251 CGAGGCGGTT GAATCACCTG AGGTCAGGAG TTCGAGACCA GCCTGACCAA  
 10301 TATCGTGAAC CTCCATCTCT ACTAAAAATA CAAAAATTAG CCAGGTGTGG  
 10351 TGGCGGGTGC TTGTAGTCCC AGTACTTGG GAGGCTGAGG CAGGTGAATT  
 10401 ACTTGAACCT GGGAGGTGGA GGTTCGAATG AGCCAAGATT GCACCACTGC  
 10451 ACTCCAGTGA CAGAGCGAGA CTCCATCTCA AAAAAAAAAA AAAAAAAGT  
 10501 TGGGAAAAGG CCAGGTGCAG TGGCTCCACG CCTGTAATCC CAACACTTTA  
 10551 AGAGGCTGAG GTGGGAGAAT CCTTTGAGCC CAGGAGTTCG AGACCAGCCT  
 10601 GGGCATTGTC CCAAGACCTT GTCTTTACAA AAAATTAGCC GGGTGTGGTG  
 10651 GCATAGCTCT GTGGTCCCAG CTATTGCGGA GGCTGAGGCA GGGAGATTGC  
 10701 TTGAGCCTAG GAGTCTAGGG CTGTAGTGAG CTGTGATCAC GTCACGTAC  
 10751 TCTAGCCTGG GCAACAGAGC AAGACTCTGT CTCCAAAAA GAAAAATAAG  
 10801 TTGGGAAAGG CTCACCTACT TCATCAGATG AGAACAAAGA CATGTTTGAA  
 10851 GTGTGAGGCC GAAGCCTGGA GAACGCTATG CGCCAGGAA ATGCAGGGCA  
 10901 GCAGAGACTC AAGATGCCAG CGCCTGTTCT GGAGGCCAG ATGGGCCCTG  
 10951 CAATGCCAC TCACCTGCC CTCCCTCTTG CCCCAGACAT CACCGGCCCC  
 11001 ATCATCCTGC AGACGTACCG CGCCATTGCC GACTACGAGA AGACCTCGGG  
 11051 CTCCGAGATG GCTCTGTCCA CGGGGACGT GGTGGAGGTC GTGGAGAAGA  
 11101 GCGAGAGCGG TCAGACCTCC CACCTTACGG GGCTCCTTCC CCTGGTGCTC  
 11151 AGGAACCCAC AGCCACAAGC CCCTGCCAA GGCTCAGGCA GCCTGGCCCC  
 11201 TGGGAGGACT CCAGCTCTGT TAGGGGCCCT AAATGTCCTC CCCCACTGT  
 11251 GGGTCGCCCT CTCTCTTAGT GTGCACCCTG TGGTGGCTGT GGGCATCTGT  
 11301 GCATGGCAGG CCGGGGCGGG GCATGTCTGC GTGTCTGTG TGGATGGGTA  
 11351 TGGGACCGTC TGTTCATTAT GAAGTGGGCT CAGAGCTGTG ATTCTGTGAG  
 11401 CATGTGTGCA TGACCTCATT GTCCAGTGTG GTGAAGGTGA  
 11451 CATTTCAAA TCTGAGCATT GGACATCAGT GTGTCTGTG CCCTGTGTCC  
 11501 TCACCATCCC TGATGGCTGC AGGAGGCCG TGGGCCCTGC CCCTCAGTCA  
 11551 CATTCCGCA CCTCTGGCAC AGGTTGGTGG TTCTGTCAGA TGAAAGCAAA  
 11601 GCGAGGCTGG ATCCCAAGC CTTCCTCGA GCCCCTGGAC AGTCCTGACG  
 11651 AGACGGAAGA CCCTGAGCCC AACTATGCAG GTGCCCCCTG CCCTCCGAGG  
 11701 CTGTAGGGGT GTGGGAGAAA GGGGCAGGCA ATATTGAGTG  
 11751 ACTGCTTTGG AGTCTGGGCT GGTGCTGGC TTGGCAGAAA AGTCAGGGCT  
 11801 AAGATCTCAT CGGCTCTGGC TTGGGGGCC TGGCAGGTTG TGATGCCCTT  
 11851 GGTCTGGACA GGAACCAAG AGGAGGAGCA GACGACTCGG GAGAGTGGGA  
 11901 GGCCAGTGGT GTCTGTGGAT ATGTGGCCAG GTTCAGTGGG AAGCTGAAGG  
 11951 ATGAGCAGAC CTTAGGCTCA GGAAGGAGGG CTGCCTGGAA GTGGGGGCAT  
 12001 CATCACTGAC CAGAAAGGGA AAACCTGGCAG TGCCAGGGCT GGATGGGGCC  
 12051 TGCATTGAGC TTGAAAAAAA CTATAATAGA ATTGGTTACC ATTTTATTTT  
 12101 ATTATTTATT TATTTATTTT ACTTTTTTGA GATAGAGTCT CACTCCCTTG  
 12151 CTAAGGCTGG AGTGCGGTGG TGCTATCTCA GCTCACTGCA ACCTCTGCCT  
 12201 CCCAGGATCA AGTGATTCTC CAGCCTCAGC CTCCCAGGT AGCTGGGATT  
 12251 ACAAGCATGC ACCACCATGC CTGGATAATT TTTGTATTTT TAGTTGAGAC  
 12301 GGGGTTTCAC CAGGTTGGCC AGACTGGTCT CGAACTTCTG ACCTCAGGTG  
 12351 ATCTGCCTGC CTCGGCCTCC CAAAGTGCTG GAATTACAGA TGTGAGCCAC  
 12401 TGTCCCTGGC CTGGTTACCC ACATTTTAAA ATGGAGTGAT TTCACCTTTT  
 12451 TATGTGGATT TACAGCTTGT TTTTTTTTTT TTTTGTGAGC AAAGTCTGGC  
 12501 TCTGTACCC AGGCTGGAGT GCAGTAATGC AATCTCAGCT CACTGCAACC  
 12551 TTAGCCTCCT GGGTTCAAGC AATTCTCCTG CCTCAGCCAC CTGAGTAGCC  
 12601 TGGGGTTTACA GGCATGCACC ACCACGCCAG GCTAATTTTT TGTATTTTTA  
 12651 GTAGAGATGG GGTTCGCCA TGTGGCCAG GCTGGTCTCG AACTCCTGAC  
 12701 CTCAGGTGAT CCGCCCGCCT TGGCCTCCCA AAGTGCTAGG ATTACAGGTG  
 12751 GGAACCACCT TGCCAGCCT GTGGCTATCG TTTAAACACT GGAAGGCCCT  
 12801 GCAGCCCCCA GGCCGACAGT TAGCTGCAGC TGAGCAGTTC CCAGTGCCAG  
 12851 GTAGACGGAT GCTCCACCCA CCTACTCATG GCTGATCTCT TGTATAGTG  
 12901 AAGTGTCTGG ACAGACCTTC ATCGTTATGG GATCTCTGGT CCCCAGAGTG  
 12951 GGTGGCAATG AATGGGAGTG GACAAGCTCA CCTGGGTGTA GGGGGCAGAG

FIGURE 3D



13001	GGCCGAAGTC	CAGAGTGTAC	CCCCAGAGTG	GGTGCCAGCA	GGAGCTTGCC
13051	GAGGGATCTG	GGATGGAGCA	GGAGGGTGGA	GGGAGGAGAC	CCAGAAGAGG
13101	GGGAAGTGTG	GGCCCTGGGT	GGGTCTGGAG	TGCCTGGAGG	AAGCCCAGGC
13151	GCAGAGAGGA	GAAGATGGGA	TGGGTGGCGA	GCCCCAGGCT	GGGCCGACCT
13201	CACACTGTGC	TCTGTGCCCC	TGCCGTGGAC	CAGGTGAGCC	ATACGTGCCC
13251	ATCAAGGCCT	ACACTGCTGT	GGAGGGGGAC	GAGGTGTCCC	TGCTCGAGGG
13301	TGAAGCTGTT	GAGGTAATTC	ACAAGCTCCT	GGACGGCTGG	TGGGTGATCA
13351	GGTAGGAGGG	CCCCTCTCCA	TCCAGAGCAC	CCATCTGAGT	CAGCCCCAGC
13401	CAGGACGGGG	TGTTTAGGGA	TCTGGGGTGA	CTTGTCCTTG	GGACTCTGGG
13451	TAAGCCACTG	CCCCTCTCTG	GGCTTAGTTT	CCATCTCAGT	AGCAGGGAGG
13501	GATGAGCCCA	CCCTTGCTTG	TCTTGTGGGG	ATCCAATGTC	CTTGTCGAAG
13551	TGGGTGCATT	TCTCCTTTGT	GATTTAGGGT	CTCTTCCCAA	CCATCTATTA
13601	TTATTCTTTC	TCTGGCAACA	TGGTGAATCG	TTGTATAAAT	AATTACATTC
13651	CTAGCTAGGC	GCAATGGCCC	AGGCCTGTAA	TCCCAGCACT	TTGGGAGCCC
13701	AGGACAGGAC	GATCAGTGTA	GGTCAGGAGT	TCGAGACCAC	CCTGGCCAAC
13751	ATGGCAAAAC	CCTATCTCTA	CTAAAAACAC	AAACATGAGC	CGGGTGTGTG
13801	GGTGGGAGCC	TGTAATCCCA	GCTACTCGGG	AGTCTGAGAC	AAGAGAATCA
13851	CTTCAACCCG	GGAGGCGGAG	GTTGCAGTGA	GCCAAGATCG	CGCCATTGCA
13901	CTCCAGCCTG	GGCAACGAGA	GCGAAACTCC	GTCTCAAAAA	AAAAAAAAAAAA
13951	AAAAAAGATT	ACTTTCTTTT	TATCATTCCT	TTATCTTTTA	AAGCTTTCTT
14001	GCAGTCAGGT	GCAGTGTCTC	ATGCCTGTAA	TCCCAACACT	TTGGGAAGCT
14051	GAGGTGGGAG	GATGACTCAA	GGCTACAAGT	TCAAGACCAA	CCTGGGCAAT
14101	GTAGGGAGAC	CTCTGTCTCT	ACAAAAAATA	TTAAAAAATA	GCTGGATGTG
14151	GACGACACAC	CCTGTAGCCC	CAGCTACTCA	GGAGGCTGAG	GTGAAAGGAT
14201	CACCTTGACCC	CAGGAGTTGG	AGGCAGCAGT	GAGCTATGAC	TGCACCACTG
14251	CACCCAGGCC	TGGGTGATGG	AGCAAGACCC	TGTCTCAAAA	AAAAAAAAAAAA
14301	AAAAAAGCT	TCCATTGCAA	TTCCCCTCTG	TTTATCCTCC	AAATGAATGC
14351	AGAAATACTA	ATTATCTTTT	TTCTGGTTCT	GGGGAACACA	GAATTCTAGC
14401	GGCTTGTGGA	GCCATTTCCT	TGGAGCCATG	GGGCCTCCCA	GGTCTTTTCC
14451	TGTGTCTTCA	TTTTTTACGA	ATTTTTTCAT	TTTTTGAGAC	AGGATCTTGC
14501	TCTGACTCCC	AGGCTGAGGC	ACAATCATCG	CTCACTCAAG	CGATCCTCCC
14551	ACCTCAGGCT	CCCACGTAGC	TGGGACTACA	GGTGAGCACC	ACCACATCTG
14601	GCTAATGTTT	TTTAATTTTT	TTGTAGGGGT	GGGGTCTCAC	TATGGTGCCA
14651	AGACTAGTCT	TAAACTCCTG	GCCTCAAGAG	TTCTCTCTGC	CTTGGCCTCC
14701	CAAAGCACTG	GGATTACAGG	AATGAGCCTC	CATGCTGGGC	CTTTGCTGGC
14751	GTCTTCAGAG	CCCTAGGTCA	CAGGGCCAGC	CTGGCGCCCT	GCCGCAAGCT
14801	TATCTTAAAG	CTGGGACCAC	AACATGCATA	CCTGCAGCCG	GGCCCCGGGG
14851	CAGAGGGCTT	TGAGGCAGCA	TTTCTCAGCC	TTTTAGACAC	ACACTCTGTT
14901	AACCCCATC	CTGTGTCTCT	GATAATCTTC	TTGTGATCCT	CCCACCAGCC
14951	AAGAATTGGG	TTTTATGTGA	ACCTTGATTT	ATGCAAGATT	TTCTTTTGTG
15001	TTTTTTTTTCA	CTCCCAAATA	TAATATTGAG	AATAGAAAGA	AAGTCTTTTC
15051	AACAAATGGT	GCTGGAACAG	ATGGATTTC	ATACTGGAAA	AAAAAAAAAAAA
15101	AGAGCAAAAA	ACAAACCTAG	ACCCCTTCCT	CACACTGTAC	ACATATGTTT
15151	ACTTCAGATG	GATCAGAGGT	TTATCCCAGA	GTAAACCTG	AAACTAAAAA
15201	CCATTTGGGG	CTGGACAGGG	AGCTCACGCC	TGTAATCTCA	GCATTTTGGG
15251	AGGCTGAGGC	AGGTGGATCA	CTTGATGTCA	GGAGTTTGAG	ACCAGCCATG
15301	ACCAATATGG	TGAAATCCTG	TCTCTACTAA	AAAAATACAA	AATTAACCAA
15351	GTGTGGTGGT	GCATGCCTGT	AATCCCAGCT	ACTTGGGAAG	CTGAGACAGG
15401	AGAATTGCTT	GAACCTGGGA	AGCAGAGGTT	GCAATGAGTC	GACATCATGC
15451	CATTGCACTC	CAGCCTAGGC	AACAAGAGCA	AAACTCTGTC	TTGGGGTTGG
15501	GTGGGGGAAA	AGCATTTGGA	AGAAAGCATA	GAATTTGGTG	GCTTGGAGGT
15551	AGGCAAGGTT	TCGTAGGAGA	CAGAAGGCAG	TTAACATAAA	AGAAAAATTG
15601	GCAAATATAA	TCCTGCCAGT	GTCTTCTTTT	TTCTTTAATT	TTTTCGGGAG
15651	GTAGAGATAG	GGGTCTTGCT	ATGTTACCCA	GGCTGATCTC	CAACTCCTGG
15701	CCTCAAGCGA	TCCTCCCAAC	TAGATCCCTC	AAAGTACTGG	GATTACAGGC
15751	GTGAGCGACC	GTGCCCTGCC	CATTCTTGCC	AATGTCTTAT	AGCAAATACC
15801	TGTCCCCTCG	GGTGACCTGG	ATCTGCTAAC	CTCCACCCCT	GCCTAGACTG
15851	TGGAAGGATT	GCTGGAAGGG	TCTCAGTTGC	ACAGACCAGG	AAACTGAGGC
15901	CCACAGAGGC	AGGTGTCCGG	TTGTTTGCAA	CCTCTCAGCC	TGTGCTAACC
15951	CCAATTGTTT	AGAGAGAGCC	CTGAAACCTT	CTCCTCTGGG	CGCCCCCAGG
16001	TGACTGCCCC	AGCCTCAAGG	GCTGCCTCTG	TTGCAGGAAA	GACGACGTCA
16051	CAGGCTACTT	CCCGTCCATG	TACCTGCAAA	AGTCAGGGCA	AGACGTGTCC
16101	CAGGCCCAAC	GCCAGATCAA	GCGGGGGGCG	CCGCCCCGCA	GGTAAGCGGG
16151	GGTCCCCGGG	GCTGGGCGGG	GTGAGCGGGG	GCGCACCACG	GGTTCGCTCT
16201	GTCTAGGCCA	TAGCTTGCCA	GTGCCGGGGC	GGGGGCTCTC	AGCCTGGCAG

FIGURE 3E

```

16251 GAGAGGCAGG ACCCTCACGG GGGAAAGGGG CTGGACGCGC CTGGCCGCGG
16301 TGTGGGGCTG GCACGGGGGC GGAAGGAAAG CGGCGATGCC CGGGGGCTTT
16351 GGGGATGGGC AGTCCAGGGG GGTCCTCCCG AGAGGGGGAC GACAGACCGA
16401 AGGCTCGTGA GGGGCGTGA AAACCGCCCA GGCTCTGCTG CAGGGCAAGG
16451 GTCCTTGTCG TGACGGGGGC AGCCGCTCTT TGTCCCGCCG GGGTCGTGCA
16501 GACTACCGGC CCCCTACTGC CCCCCACTTC CTCGACCAG GGGTGCCCAT
16551 CTGAGTCCCT GGGGGCAGGG GCGCCCTCGG GCTTTGACGA CGCCCCGTCC
16601 CGCTGGGCCA GGTGCTCCAT CCGCAACGCG CACAGCATCC ACCAGCGGTC
16651 GCGGAAGCGC CTCAGCCAGG ACGCTATCG CCGCAACAGC GTCCGTTTTC
16701 TGCAGCAGCG ACGCCGCCAG GCGCGGCCGG GACCGCAGAG CCCCAGGAGC
16751 CCGCTCGGTG AGTGACGCGG GAGAGGGCAG GAAGGGCAAG CCCTAGGGGC
16801 GGAGTCAGCG GGAGAGGCGG GGCCAGAGGC AGGGCCAGAG TAGCGGGGCG
16851 GGACCAGAGG GCGGAATCAG AGGGAGAGGC GGGGACTGGA GGCGGGGGCA
16901 GAGGAGGAGC CAGCGCTAGG GGGCGGAGCG ATCCCTAAGA GGCGGAGTCA
16951 GAGGGAGAGG CACAAGCGGG AGGCGAGGCC AGAGCGCGGA GCAGGAGTTG
17001 GAGACCGCGG CCGGGCAGGG CCAGAGAGCG CTGTGGGCGG GGCCAGTGTG
17051 CGGGGCGGGG CGTCTGACTC GGCCCGGCTC TCTGCCCGCA GAGGAGGAGC
17101 GGCAGACGCA GCGCTCTAAA CCGCAGCCGG CGGTGCCCCC GCGGCCGAGC
17151 GCCGACCTCA TCCTGAACCG CTGCAGCGAG AGCACCAGC GGAAGCTGGC
17201 GTCTGCCGTC TGAGGCTGGA GCGCAGTCCC CAGTAGCGT CTCGGCCCTT
17251 GCCGCCCGCT GCCTGTATAT ACGTGTCTA TAGAGCCTGG CGTCTGGAGC
17301 CCGAGGGCGT CCGGACCCCG TGTCAGCGCG GGCTCCCGCC ACCCTCAATA
17351 AATGTTGCTT CGAGTGAGCC GAGGCTCTGC AGGAATGCAG GGAGGGCCGG
17401 GCTCCGCCCC AGGGTTATTT TCTAAGTTGA GGACAGGGAG GTTGTGAGTT
17451 CTGNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18051 NNNNNNNNNN NNNNNNTAAA AATTAGCTGG GCGTGGTGGC ATGCATCCAC
18101 AATCCAGCT ACTGGGAGG CTGAGGCATG AGAATCGCTT GAACCGGGGA
18151 GGCAGATGTT GCAGTGAGCC GAGACGGCGC CACTGCACTC CAGCCTGGAC
18201 TACAGAGCGA GACTCTATCT CAAAAA AAAA AAGTAACTTA
18251 GGTGCAGGGT GTCCTCTGTT ATTCACTGAG ACCGTGCCCC GGTTATGAGG
18301 TTGTACCAGA AAGCAAGTAT TCACTATGCA CACTATTAC CGCTCACCTT
18351 AGCATTGAAG CCAGCCTGTA GCCTGAAAGC CTTTGCTTTG AGGGCAGGTC
18401 TTTCCCAAAA ATGCAGACAC GAAGGTGCAA AGTGAAGCTG CCAGTCTTGC
18451 AAAAGATGTA AATTGTACG AAGGCCACGA GTGGCAGGGA GAGCTGTCCC
18501 ACATTTGCGG AAGTGGCTAT GTGAGGACGG GGGAGCGCGG TCCCTTAGAG
18551 ATAAGAGACA ATCATAAGGG GAGATATCAG AGAAAATCGT AAGGGGAGCA
18601 GATGGTTGTC AAGAGAATAG GCTGACCATC GAAGGACTGG CAGAAGCTTT
18651 CAGAAAACCA CTGGACGGCT GGGCACAGTG GCTTAGGCCT GTAATCCAG
18701 CACTTTGGGA GGCTGACGCA GGTGAATCAC TTGAGGTCAG GAGTTCCAGA
18751 CCAGCTGGC CAACATGGTG AAACCCATC TCTACAGAAA ATATAAAAT
18801 TAGCCAGGCG TGGTGGCACA AGCCTAGAAT CCCAGCTACT TGGGAGGCTG
18851 AGG

```

(SEQ ID NO: 3)

#### FEATURES:

Start: 2038  
 Exon: 2038-2109  
 Intron: 2110-4800  
 Exon: 4801-4805  
 Intron: 4806-5298  
 Exon: 5299-5377  
 Intron: 5378-7131  
 Exon: 7132-7207  
 Intron: 7208-7306

FIGURE 3F

Exon: 7307-7472  
 Intron: 7473-8829  
 Exon: 8830-8885  
 Intron: 8886-10986  
 Exon: 10987-11109  
 Intron: 11110-11572  
 Exon: 11573-11680  
 Intron: 11681-13233  
 Exon: 13234-13339  
 Intron: 13340-16036  
 Exon: 16037-16141  
 Intron: 16142-16611  
 Exon: 16612-16757  
 Intron: 16758-17091  
 Exon: 17092-17210  
 Stop: 17211

SNPs:

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
139	T	C	Beyond ORF(5')			
262	T	C	Beyond ORF(5')			
43	A	G	Beyond ORF(5')			
344	G	A	Beyond ORF(5')			
721	A	G	Beyond ORF(5')			
1038	A	-	Beyond ORF(5')			
1468	A	G T	Beyond ORF(5')			
2758	G	T	Intron			
2890	G	A	Intron			
3304	G	A T	Intron			
3896	C	T	Intron			
3906	-	C G	Intron			
3907	-	T	Intron			
3911	-	T	Intron			
3932	-	A	Intron			
3934	-	T	Intron			
3949	C	T	Intron			
3994	A	G	Intron			
6272	C	T	Intron			
6427	T	C	Intron			
6694	T	C	Intron			
7741	C	T	Intron			
8294	A	G	Intron			
9313	A	- G	Intron			
10838	G	A	Intron			
11093	G	A	Exon	187	E	E
11195	T	G	Intron			
11213	G	A	Intron			
11263	C	G	Intron			
13707	G	A	Intron			
14629	G	A	Intron			
14698	T	A	Intron			
16095	C	T G	Exon	284	P	S A
16266	C	T	Intron			
16629	C	T	Exon	305	P	L
16642	C	T	Exon	309	H	H
18537	C	T	Beyond ORF(3')			
18589	G	A	Beyond ORF(3')			
18720	G	A	Beyond ORF(3')			
18782	C	T	Beyond ORF(3')			
18841	C	T	Beyond ORF(3')			

FIGURE 3G

Context:

DNA  
Position

139 TACTAAAAATACAAAATTAGCCAGGCGTGGTGGCGCACACCTGTAATCCCAGCTACTTGG  
GAAGCTGAGGCAGGAGAATCGCTTGAACCTGGAAGGCAGAGGTTGCAGTGAGCCGAGATT  
GTGCCACTGCACTCCAGC  
[T,C]  
TGGGCAACAAGAGCGAAACTTCGCTTCAAACAAATAAATTAACGCCCAGCATGTCTTGGC  
TTTCATCTGCCAGACCTCAACCCCTACCCCCAGGAGATCAGGTCCGACCATGAGCTGAC  
CCTGGACTCAGGCAAGGGTGAGTTGGTGCAGCCCTGGCCTGCTGGGAGGCACAGGCTGCA  
GCAGGCTGCCTGGGGCTGAGGCCCGCCACTCATGAACTCATGACCTTGAATGAGCTCCAA  
AAGCTCTGGGCCTCCAGGCTCTAGGGGGAGTGGGAGAGAGAGGCTCAGCCTGTCCCTG

262 TACTAAAAATACAAAATTAGCCAGGCGTGGTGGCGCACACCTGTAATCCCAGCTACTTGG  
GAAGCTGAGGCAGGAGAATCGCTTGAACCTGGAAGGCAGAGGTTGCAGTGAGCCGAGATT  
GTGCCACTGCACTCCAGCCTGGGCAACAAGAGCGAAACTTCGCTTCAAACAAATAAATTA  
ACGCCCAGCATGTCTTGGCTTTCATCTGCCAGACCTCAACCCCTACCCCCAGGAGATCAG  
GTCCGACCATGAGCTGACCC  
[T,C]  
GGACTCAGGCAAGGGTGAGTTGGTGCAGCCCTGGCCTGCTGGGAGGCACAGGCTGCAGCA  
GGCTGCCTGGGGCTGAGGCCCGCCACTCATGAACTCATGACCTTGAATGAGCTCCAAAAG  
CTCTGGGCCTCCCAGGCTCTAGGGGGAGTGGGAGAGAGAGGCTCAGCCTGTCCCTGGGC  
ATGCTGCCCCCTCCTCACCTCTTGTCCCAAATCCCCTTCTTGGCAAAGCTGACAGTCTT  
AATATCACTCTGGAGAAAAGTGAATCAGCCCTAAGGAACAATTCAATGA

43 TACTAAAAATACAAAATTAGCCAGGCGTGGTGGCGCACACCT  
[A,G]  
TAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAATCGCTTGAACCTGGAAGGCAGAGGT  
TGCAGTGAGCCGAGATTGTGCCACTGCACTCCAGCCTGGGCAACAAGAGCGAAACTTCGC  
TTCAAACAAATAAATTAACGCCCAGCATGTCTTGGCTTTCATCTGCCAGACCTCAACCCCT  
CACCCCCAGGAGATCAGGTCCGACCATGAGCTGACCTGGACTCAGGCAAGGGTGAGTT  
GGTGCAGCCCTGGCCTGCTGGGAGGCACAGGCTGCAGCAGGCTGCCTGGGGCTGAGGCC

344 TAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAATCGCTTGAACCTGGAAGGCAGAGGT  
TGCAGTGAGCCGAGATTGTGCCACTGCACTCCAGCCTGGGCAACAAGAGCGAAACTTCGC  
TTCAAACAAATAAATTAACGCCCAGCATGTCTTGGCTTTCATCTGCCAGACCTCAACCCCT  
CACCCCCAGGAGATCAGGTCCGACCATGAGCTGACCTGGACTCAGGCAAGGGTGAGTT  
GGTGCAGCCCTGGCCTGCTGGGAGGCACAGGCTGCAGCAGGCTGCCTGGGGCTGAGGCC  
[G,A]  
CCACTCATGAACTCATGACCTTGAATGAGCTCCAAAAGCTCTGGGCCTCCCAGGCTCTAG  
GGGAGTGGGAGAGAGAGGCTCAGCCTGTCCCTGGGCATGCTGCCCCCTCCTCACCTCT  
TTGTCCCAAATCCCCTTCTTGGCAAAGCTGACAGTCTTAATATCACTCTGGAGAAAAGT  
AGTCAGCCCTAAGGAACAATTCAATGAACCATTTGCTTACTTGAGGATTGGAATCAAGT  
CTCACTCAAAGTCTGTGCCATTTTCGTCCCAGCTGTCACTGGCCCTCATCCACACACCC

721 AGGCCTCAGCCTGTCCCTGGGCATGCTGCCCCCTCCTCACCTCTTGTCCCAAATCCCCT  
TCCTGGCAAAGCTGACAGTCTTAATATCACTCTGGAGAAAAGTGAATCAGCCCTAAGGAA  
CAATTCAATGAACCATTTGCTTACTTGAGGATTGGAACCTCAAGTCTCACTCAAAGTCTGT  
GCCATTTTCGTCCCAGCTGTCACTGGCCCTCATCCACACACCCCAAGGATGAGCATCTA  
ACGCTTGATGCACACTCCCATGCCCCGCTTCAATCACTATTCAATCATTCAATCACTC  
[A,G]  
TTCATTGACTCATTCAATCATTCACTCACTCATTCAATCACTCAGTGAATGTTGCAGTCA  
CGATCCAAATATTTATGGCCTCTGTGTGCCAGGCACTAGATGGAGGGGCTGGGGCTAGAG  
CCCCTGATAACCCGGTCATGCCCTAGCTTTCCTGGGACACACATTGTGGTAAGGGGAGAC  
TAAAAAATTAAGTCAGGCCAGGCACGGTGGCTCATGCCTGAATCCCAGCACTTTGGGAG  
GCCGAGGCAGTGAATTACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCCAACATGGAGA

1038 TTCATTCACTCACTCATTCAATCACTCAGTGAATGTTGCAGTCACGATCCAAATATTTAT  
GGCCTCTGTGTGCCAGGCACTAGATGGAGGGGCTGGGGCTAGAGCCCTGATAACCCGGT  
CATGCCCTAGCTTTCCTGGGACACACATTGTGGTAAGGGGAGACTAAAAAATTAAGTCA  
GGCCAGGCACGGTGGCTCATGCCTGAATCCCAGCACTTTGGGAGGCCGAGGCAGTGAAT  
TACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCCAACATGGAGAAACCCAGTCTCTAATT  
[A,-]

FIGURE 3H

AAAAAAAAAAAAAATTAGCCAGGTGTGGTGGCACATGCCTGTAATCCCAGCTACTCAGG  
 AGACTAACGCAAGAGAATTGCTTGAACCCAGGAGGCAGAGGTGCGGTGAGCCGAGATCG  
 CGCCATTGCACTCCAGCCTGGGAAACAAGAGCGAGACTCCATCTCAAAAAAAAAAAGT  
 GGGAGGCAGAGGCAGGAGGATCACTAGAGGCCAGTAGTTTGAGACCATCCTGGGCAACAT  
 AGCAGGACCCTGTCTGTACAAAAAATTAAAAAATTTAACCGGGCATGGTGGCACACA

1468 ACTCCAGCCTGGGAAACAAGAGCGAGACTCCATCTCAAAAAAAAAAAGTGGGAGGCAG  
 AGGCAGGAGGATCACTAGAGGCCAGTAGTTTGAGACCATCCTGGGCAACATAGCAGGACC  
 CTGTCTGTACAAAAAATTAAAAAATTTAACCGGGCATGGTGGCACACACCCGTAGTC  
 CCAGCTACTCCAGAGGCTGAGGCAGGAGGATCGCTGGAGCCAGGAGTTGGAGGCTGCAG  
 TGAAGTGTGATCCCACTGCACCTAAGCCTGGATAACAAAGCAAGACCCTGTCTCAA  
 [A,G,T]  
 AACAATAGCAATAATAATAAGAAAAATTAATGCAATTTGCGATGCATCAGTGATAAGT  
 GCTCTGCAGAAAAAGGAGGCAGGAAGAGGCTGAGAAAGGTATGAGTTTGCTATGCAATG  
 TGAAGTTATCAAGGAAGGCTTCTCGGAAGAGGTGACATTTGAGCAGAGAAATGGAGGAGA  
 GTTATGGAGGGAAGATGGTGAATGGGGGAACATGGTCAAGACCAGGAATATGGTCAAGG  
 GGGAAAGATGGTCAAGGGGACGCAGCAATGCAAAGGCCCTGAGGCAGGAGCAGCTTGA

2758 TCAGAGAGTGGCCCTCCCAAGAGGGAGGCCTGGAACTAAAGCTCTCTCTCCCCAGCT  
 GCCTGTAGTGTGAGTGTAGAGTCTTATCTCTCCAGTAGGGTGACACCATGACAGGGGCCA  
 ATAGAGTCTCTCCATCTGTCCCCAAGGAGGCTGGACAAATGCCTGCTCAGACACACAAGT  
 CCAGTCTGCTCCCTAATCCCATAGGAAGGCCAGGGAGGAACACTACATTTAGGAAATTTGAG  
 CTTGTATGGAAACATTTAGTCTTATGTGCAAGACCTTTCTCTTTTGTATTCTTTTGT  
 [G,T]  
 TTTTGAGACAGAGTCTTGATCTGTTGCCAGGCCAGAGTGCAGTGGCAGCATCTCAGCTC  
 ACTGCAACCTCCGCCTTCCAGGTTCAACTGGTTCTCTGCCTCAGCCTCCAGAGTAGTTG  
 GGATTACAGGTGCCACCACCACGCCTGGCTAATTTTGTATTTTGTAGTAGACAGGGT  
 TTCACCATGTTGGCCAGACTGGTCTCAAACTCCTGACCTCAAGTGATCCACCACCTGGG  
 CCTCCAAAGTGCTGGGATTACAGGCATGAGCCACCGTGCCTGGCCTGTTTTTTTGAAT

2890 CATCTGTCCCCAAGGAGGCTGGACAAATGCCTGCTCAGACACACAAGTCCACTGGGTCCC  
 CTAATCCCATAGGAAGGCCAGGGAGGAACTACATTTAGGAAATTTAGAGCTTGTATGGAAC  
 ATTTAGTCTATGTGCAAGACCTTTCTCTTTTGTATTCTTTTGTGTTTGTAGACAG  
 AGTCTTGATCTGTTGCCAGGCCAGAGTGCAGTGGCAGCATCTCAGCTCACTGCAACCTC  
 CGCCTTCCAGGTTCAACTGGTTCTCTGCCTCAGCCTCCAGAGTAGTTGGGATTACAGGT  
 [G,A]  
 CCCACCACCACGCCTGGCTAATTTTGTATTTTGTAGTAGACAGGGTTTACCATGTTG  
 GCCAGACTGGTCTCAAACTCCTGACCTCAAGTGATCCACCACCTGGGCCTCCCAAAGTG  
 CTGGGATTACAGGCATGAGCCACCGTGCCTGGCCTGTTTTTTGAAATGAGGTCTGGAGT  
 GCAGTGGTCCGATCATAGTTCAGTGCAGCCTCAACCTCCAGGCCCAAGTGATCCTCCTG  
 CCTCAGCCCTTGAGTAGCTGGGGCTACAGGCGCACACCACCATGCCTGGCTAGTTTTTA

3304 CAAAGTGCTGGGATTACAGGCATGAGCCACCGTGCCTGGCCTGTTTTTTGAAATGAGGT  
 CTGGAGTGCAGTGGTGCATCATAGTTCAGTGCAGCCTCAACCTCCAGGCCCAAGTGAT  
 CCTCCTGCCTCAGCCCTTGAGTAGCTGGGGCTACAGGCGCACACCACCATGCCTGGCTA  
 GTTTTTAAAAATTTTGTGGAGATGAGGTTTCACTATGTTGTCCAGGCTAATCTTGAACCT  
 CTCGGCTTAAGCAACCTCTGGTCTCAGCCTCCACAGTGCTAGGATTACAAGCGTGAGC  
 [G,A,T]  
 ACCGTGCCTAGTCACTTTTCTCCTTTCTTTGTAACCTTCAGTTTTGAAATTTCAAATTT  
 ACAGAAAGGCTACTGGGTGTCAAAACGGTACCAGTCACTCCAATAGTCTTCACTCACCT  
 TCATCCACACCTCTTTCTGGGGATATTTCTGAATTATTTGAGAGTGAGTTGAAGACG  
 TGTCTTTTACCTCTAAATACTAGTTGTTGGGCATTTCTTAAATCAAGGCATTCTCTTA  
 CATAATCAACACACGTGTCAAAATCAGGAAATTAACATGGACAAAACACCATTTATCCA

3896 CATTATCCACCCACAGACTTTACTGAGGTTTCCCGATTATCCTGCTTGTCTCTGCACT  
 GAAAACTTTTTTCAAGTCTAGGATCCAGTCAAGGATCAATGTATAGCCTTTAACCTTCT  
 TTAATCTGGATCAGTCTTTTTCTTTTTCTTTTTCTTTTTTGGACACGGAATCTCACTC  
 TGTGCCCAGACTGGAGTGCAGTGGTGAATCTCGGCTATTGCAACCTCTGCCTCCTGGG  
 TTCAAGAGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGAATACAGGTGCGGCCACCA  
 [C,T]  
 GCCCAGCTCGTTTTTGGTAGAGACAGGGTTTTGCCATTGATTCTGGATCAGTCTTTTTT  
 TTTTTTTATGAGATGGAGTCTTACTCTGTCAACCCAGGCTGGAGTGCAATGGCACAATCT  
 CCACTCACTGCATCCTCCGCTCCAGGTTCAAGCAATTTCTGTCCTCAGCCTCCCGAG  
 TAGCTGGGATTACAGGCATGCGCCACCATGCCCCGCTACTTTTTGTATTTTGTAGTAGA  
 CAGGTTTACCATGTTAGCCAGGCTGATCTCGAACTCCTGACGTGAGGTGATCTGCCCG

FIGURE 3I

3906 CCACAGACTTTACTGAGGTTTCCCCGATTATCCTGCTTGTCTCTGCAGTGAAAACTTTT  
TTCAGGTCTAGGATCCAGTCAAGGATCAATGTCATAGCCTTTAACCTTCTTTAATCTGGA  
TCAGTCTTTTTCTTTTTCTTTTTCTTTTTTGGACACGGAATCTCACTCTGTGCGCCAGA  
CTGGAGTGCAGTGGTGAATCTCGGCTCATTGCAACCTCTGCCTCCTGGGTTCAAGAGAT  
TCTCTGCCTCAGCCTCCTGAGTAGCTGGGAATACAGGTGCGCGCCACCACGCCAGCTC  
[-,C,G]  
TTTTTGGTAGAGACAGGGTTTTGCCATTGATTCTGGATCAGTCTTTTTTTTTTTTTAT  
GAGATGGAGTCTTACTCTGTCACCCAGGCTGGAGTGCAATGGCACAATCTCCACTCACTG  
CATCCTCCGCTCCAGGTTCAAGCAATTCTCGTGCCTCAGCCTCCCGAGTAGCTGGGAT  
TACAGGCATGCGCCACCATGCCCCGCTACTTTTTGTATTTTAGTAGAGACAGGGTTTCA  
CCATGTTAGCCAGGCTGATCTCGAACTCCTGACGTCAGGTGATCTGCCCCGCTCGACCTC

3907 CACAGACTTTACTGAGGTTTCCCCGATTATCCTGCTTGTCTCTGCAGTGAAAACTTTT  
TCAGGTCTAGGATCCAGTCAAGGATCAATGTCATAGCCTTTAACCTTCTTTAATCTGGAT  
CAGTCTTTTTCTTTTTCTTTTTCTTTTTTGGACACGGAATCTCACTCTGTGCGCCAGAC  
TGGAGTGCAGTGGTGAATCTCGGCTCATTGCAACCTCTGCCTCCTGGGTTCAAGAGATT  
CTCCTGCCTCAGCCTCCTGAGTAGCTGGGAATACAGGTGCGCGCCACCACGCCAGCTCG  
[-,T]  
TTTTTGGTAGAGACAGGGTTTTGCCATTGATTCTGGATCAGTCTTTTTTTTTTTTTATG  
AGATGGAGTCTTACTCTGTCACCCAGGCTGGAGTGCAATGGCACAATCTCCACTCACTGC  
ATCCTCCGCTCCAGGTTCAAGCAATTCTCGTGCCTCAGCCTCCCGAGTAGCTGGGATT  
ACAGGCATGCGCCACCATGCCCCGCTACTTTTTGTATTTTAGTAGAGACAGGGTTTCA  
CATGTTAGCCAGGCTGATCTCGAACTCCTGACGTCAGGTGATCTGCCCCGCTCGACCTCC

3911 GACTTTACTGAGGTTTCCCCGATTATCCTGCTTGTCTCTGCAGTGAAAACTTTTTT  
GTCTAGGATCCAGTCAAGGATCAATGTCATAGCCTTTAACCTTCTTTAATCTGGATCAGT  
CTTTTTCTTTTTCTTTTTCTTTTTTGGACACGGAATCTCACTCTGTGCGCCAGACTGGA  
GTGCAGTGGTGAATCTCGGCTCATTGCAACCTCTGCCTCCTGGGTTCAAGAGATTCTCC  
TGCCTCAGCCTCCTGAGTAGCTGGGAATACAGGTGCGCGCCACCACGCCAGCTCGTTTT  
[-,T]  
GGTAGAGACAGGGTTTTGCCATTGATTCTGGATCAGTCTTTTTTTTTTTTTATGAGAT  
GGAGTCTTACTCTGTCACCCAGGCTGGAGTGCAATGGCACAATCTCCACTCACTGCATCC  
TCTCTCGTCCAGGTTCAAGCAATTCTCGTGCCTCAGCCTCCCGAGTAGCTGGGATTACAG  
GCATGCGCCACCATGCCCCGCTACTTTTTGTATTTTAGTAGAGACAGGGTTTCAACATG  
TTAGCCAGGCTGATCTCGAACTCCTGACGTCAGGTGATCTGCCCCGCTCGACCTCCAAA

3932 ATTATCCTGCTTGTCTCTGCAGTGAAAACTTTTTTTCAGGTCTAGGATCCAGTCAAGGAT  
CAATGTCATAGCCTTTAACCTTCTTTAATCTGGATCAGTCTTTTTCTTTTTCTTTTTCTT  
TTTTTGGACACGGAATCTCACTCTGTGCGCCAGACTGGAGTGCAGTGGTGAATCTCGGCT  
TCATTGCAACCTCTGCCTCCTGGGTTCAAGAGATTCTCCTGCCTCAGCCTCCTGAGTAGC  
TGGGAATACAGGTGCGCGCCACCACGCCAGCTCGTTTTTGGTAGAGACAGGGTTTTGCC  
[-,A]  
TTGATTCTGGATCAGTCTTTTTTTTTTTTTATGAGATGGAGTCTTACTCTGTCACCCA  
GGCTGGAGTGCAATGGCACAATCTCCACTCACTGCATCCTCCGCTCCAGGTTCAAGCA  
ATTCTCGTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACCATGCCCGGC  
TACTTTTTGTATTTTAGTAGAGACAGGGTTTCAACATGTTAGCCAGGCTGATCTCGAAC  
TCCTGACGTCAGGTGATCTGCCCCGCTCGACCTCCAAAGTGCTGGGATTACAGGCGTG

3934 TATCCTGCTTGTCTCTGCAGTGAAAACTTTTTTTCAGGTCTAGGATCCAGTCAAGGATCA  
ATGTCATAGCCTTTAACCTTCTTTAATCTGGATCAGTCTTTTTCTTTTTCTTTTTCTTT  
TTTTTGGACACGGAATCTCACTCTGTGCGCCAGACTGGAGTGCAGTGGTGAATCTCGGCT  
ATTGCAACCTCTGCCTCCTGGGTTCAAGAGATTCTCCTGCCTCAGCCTCCTGAGTAGCTG  
GGAATACAGGTGCGCGCCACCACGCCAGCTCGTTTTTGGTAGAGACAGGGTTTTGCCAT  
[-,T]  
GATTCTGGATCAGTCTTTTTTTTTTTTTATGAGATGGAGTCTTACTCTGTCACCCAGG  
CTGGAGTGCAATGGCACAATCTCCACTCACTGCATCCTCCGCTCCAGGTTCAAGCAAT  
TCTCGTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACCATGCCCGGCTA  
CTTTTTGTATTTTAGTAGAGACAGGGTTTCAACATGTTAGCCAGGCTGATCTCGAACTC  
CTGACGTCAGGTGATCTGCCCCGCTCGACCTCCAAAGTGCTGGGATTACAGGCGTGAGC

3949 CTGCAGTGAAAACTTTTTTTCAGGTCTAGGATCCAGTCAAGGATCAATGTCATAGCCTTTA  
ACCTTCTTTAATCTGGATCAGTCTTTTTCTTTTTCTTTTTCTTTTTTGGACACGGAAT  
CTCACTCTGTGCGCCAGACTGGAGTGCAGTGGTGAATCTCGGCTCATTGCAACCTCTGCC  
TCCTGGGTTCAAGAGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGAATACAGGTGCGC

FIGURE 3J

GCCACCACGCCAGCTCGTTTTTGGTAGAGACAGGGTTTTGCCATTGATTCTGGATCAGT  
 [C,T]  
 TTTTTTTTTTTTTATGAGATGGAGTCTTACTCTGTCACCCAGGCTGGAGTGCAATGGC  
 ACAATCTCCACTCACTGCATCCTCCGCCTCCAGGTTCAAGCAATTCTCGTCCCTCAGCC  
 TCCCAGTAGCTGGGATTACAGGCATGCGCCACCATGCCCGGCTACTTTTTGTATTTTA  
 GTAGAGACAGGGTTTACCATGTTAGCCAGGCTGATCTCGAACTCCTGACGTGAGGTGAT  
 CTGCCCGCCTCGACCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCAGCGGA

3994 ATGTCATAGCCTTTAACTTCTTTAATCTGGATCAGTCTTTTTCTTTTTCTTTTTCTTT  
 TTTTGGACACGGAATCTCACTCTGTCCGAGACTGGAGTGCAAGTGGTGAATCTCGGCTC  
 ATTGCAACCTCTGCCTCCTGGGTCAAGAGATTCTCTGCCTCAGCCTCCTGAGTAGCTG  
 GGAATACAGGTGCGCGCCACCAAGCCAGCTCGTTTTTGGTAGAGACAGGGTTTTGCCAT  
 TGATTCTGGATCAGTCTTTTTTTTTTTTTATGAGATGGAGTCTTACTCTGTCACCCAG  
 [A,G]  
 CTGGAGTGCAATGGCACAATCTCCACTCACTGCATCCTCCGCCTCCAGGTTCAAGCAAT  
 TCTCGTCCCTCAGCCTCCGAGTAGCTGGGATTACAGGCATGCGCCACCATGCCCGGCTA  
 CTTTTGTATTTTAGTAGAGACAGGGTTTACCATGTTAGCCAGGCTGATCTCGAACTC  
 CTGACGTGAGGTGATCTGCCCGCCTCGACCTCCCAAAGTGCTGGGATTACAGGCGTGAGC  
 CACCGTGCCAGCGGATTCTGGATCGGTCTTAATCAGTCTTTGTCTTTGCAACTTTGATG

6272 AAAGTAAACAGACAGGATCTCCAGAACCTTCTAGAATGGAACCATCTTGTGCTTTT  
 TGAAAAACAAAGCCAAGTTCTAGATCCCAAATAAATGCACCTGCTGGTGAACATTCTCCT  
 TGTGGTTCTCGTCCCTATGTTAGTTATTTCTCTAAATTTTACATTTGTACCTTTTAAAGA  
 ATGAGTTATCAGTTTTTTATATTTGCTTTTCTTTGAGATGGGGTCTTGTCTGTCAACC  
 CAGGCTGGGGTGAGTGGTGAATCACGGCTCACTGCAGCCTCAACCTCCAGGGCTGAAG  
 [C,T]  
 GATTCTCCCATCTCAGCCTCCCATGTTGAGATCACAGGTGTGCACCACCACACCTGGCTC  
 CTTTTCTGATTTGTTTTTTGTAGAGATGGGATTTGCTATGTTGCCCAGGCTGGTCTCT  
 AACTCCTGGACTCAAGTGATCCTCCGCCTCAGCTTCCCAAATGCTAGGATTACAGGTT  
 TGAGCCCTGCACCTGGTCAACCTGAGTTTTAAGAGGATCCCTTTGGCGACTGGATTGAG  
 GACAGACAAGAGTGACGGGGGACACAAGGAGGCCATTTTCGTTATCCAGGCTGGTAGT

6427 ATTTTACATTTGTACCTTTTAAAGATGAGTTATCAGTTTTTTATATTTGCTTTTCTTT  
 TGAGATGGGGTCTTGCTCTGTCACCCAGGCTGGGGTGAGTGGTGAATCACGGCTCACT  
 GCAGCCTCAACCTCCAGGGCTGAAGCGATTCTCCCATCTCAGCCTCCCATGTTGAGATCA  
 CAGGTGTGCACCACCACACCTGGCTCCTTTTCTGATTTGTTTTTTGTAGAGATGGGATT  
 TCGCTATGTTGCCCAGGCTGGTCTCTAACTCCTGGACTCAAGTGATCCTCCCGCCTCAGC  
 [T,C]  
 TCCCAAATTGCTAGGATTACAGGTTTGAGCCCCTGCACCTGGTCAACCTGAGTTTTAAGA  
 GGATCCCTTTTGGCGACTGGATTGAGGACAGACAAGAGTGAGCGGGGACACAAGGAGGCC  
 ATTTTCTGTTATCCAGGCTGGTAGTGGTAGGGCCAGGAGGGTGGGGTGGTGGGAAGCA  
 GTCAGATCCCAAAGAGATTTGGGGATTGGAAGCAAAAGGATTTGCTGGTGACTTGACAT  
 GGGAGGGAGAGAGGTGAGTGCCTCTGTTAATCAAGGAATCCAGATTGCCACCGAAATTC

6694 ACTCCTGGACTCAAGTGATCCTCCCGCCTCAGCTTCCCAAATGCTAGGATTACAGGTTT  
 GAGCCCCCTGCACCTGGTCAACCTGAGTTTTAAGAGGATCCCTTTGGCGACTGGATTGAGG  
 ACAGACAAGAGTGAGCGGGGACACAAGGAGGCCATTTTCGTTATCCAGGCTGGTAGTG  
 GCTAGGGCCAGGAGGGTGGGGTGGTGGGAAGCAGTCAGATCCCAAAGAGATTTGGGGAT  
 TGGAAAGCAAAAGGATTTGCTGGTGACTTGACATGGGAGGGAGAGAGGTGAGTGCCTCTG  
 [T,C]  
 TAATCAAGGAATCCAGATTGCCACCGAAATTTCTAGGCCCGAGATATTTAGGTAGTGTCT  
 CACTCTGTCACCCAGGATGGAGTGAGTGAGTGGCGCCATCTCGGCTCACTGTAACCTCCGCCT  
 CCCAGGTTTAAACGATTCTCCACCTCAGCCTCCTGAGTAGCTGGGATTACAGGCATGTG  
 CCACCACTCCCGGCTAATTTTTGTATTTTAGTAGAGACGGGGTTTACCACGTTGGCCA  
 GGCTGGTCTTGAACCTCTGACCTCAAGTGATCCACCCACGACAGCCTCCCAAAGTGCTGG

7741 CCTGATGACCTCAAGCTCCCAAGGACAACCAAGTGAGTGAACTTTTACCCCTGCCAGGTG  
 GGAGAGGGAAGGAGGGGTGGGACTTTCTGTGTTTTGCAGATGAGGAAACCAAGGCTCAGA  
 GAGGGAAAGCCACCTTCCAGAGCCACACAGCCAGAAAGAGGAGGCAAAATCCACCTCCG  
 GCCCTGTGACCCCGCAAGCCTCACCTTAATCTTTACACCTCAGGGCACTGGGGGAA  
 GCACTCGGGGCTGGAGGTTCAAAGTCTGGGTCTCATCTGACATTATGGCCACCTGGC  
 [C,T]  
 ATGGGACCTGGAGCCAGTCACCACTGCTCTGTAATGCAGGTTCTCCATTTCTATAATGG  
 GCAGTGAGGATCAGATGAAGCATTGGGTGCTTTCGGGAGCCCCCAGAAGGATGTGGGGT  
 TGATGCCTCTGCTAAGTGCTGAGCATGTCTGGGTCTCTGTACCCAGGACCTGTGTGG

FIGURE 3K





TTCTGTCTGGATGGGTATGGGACCGTCTGTTCAATTATGAAGTGGGCTCAGAGCTGTGATT  
 CTGTGAGCATGTGTGCATGCATGTGACCTCATTGTCCAGTGTGGTGAAGGTGACAT  
 TTCCAAATCTGAGCATTGGACATCAGTGTGTCTGTGTCCCTGTGTCTCACCATCCCTGA

11263 ACCCTGCCCTCCCTCTTGCCCCAGACATCACCGGCCCCATCATCTGCAGACGTACCGCG  
 CCATTGCCGACTACGAGAAGACCTCGGGCTCCGAGATGGCTCTGTCCACGGGGGACGTGG  
 TGGAGGTCGTGGAGAAGAGCGAGAGCGGTGAGACCTCCACCTTACGGGGCTCCTTCCCC  
 TGGTGCTCAGGAACCCACAGCCACAAGCCCCCTGCCAAGGCTCAGGCAGCCTGGCCCTG  
 GGAGGACTCCAGCTCTGTTAGGGGCCCTAAATGTCTCCCCACACTGTGGGTGCGCTTCT  
 [C, G]  
 TCTTAGTGTGCACCTGTGGTGGCTGTGGGCATCTGTGCATGGCAGGCCGGGGCGGGGCA  
 TGTCTCGTGTCTGTCTGGATGGGTATGGGACCGTCTGTTCAATTATGAAGTGGGCTCAG  
 AGCTGTGATTCTGTGAGCATGTGTGCATGCATGTGACCTCATTGTCCAGTGTGGTG  
 AAGGTGACATTTCCAAATCTGAGCATTGGACATCAGTGTGTCTGTGTCCCTGTGTCTCA  
 CCATCCCTGATGGTGCAGGAGCGCTGGGCCCTGCCCTCAGTCACATTTCCCGACCT

13707 GGGGTGTTTAGGGATCTGGGGTGACTTGTCCCTGGGACTCTGGTAAGCCACTGCCCTC  
 TCTGGGCTTAGTITCCATCTCAGTAGCAGGGAGGGATGAGCCACCCTTGCTGTCTTGT  
 GGGGATCCAATGTCTTGTCCAAGTGGGTGATTCTCCTTTGTGATTTAGGGTCTCTTC  
 CCAACCATCTATTATTCTCTCTGGCAACATGGTGAACGTGTGTATAAATAATTAC  
 ATTCCTAGCTAGGCGCAATGGCCAGGCTGTAATCCAGCACTTTGGGAGCCCAGGACA  
 [G, A]  
 GACGATCAGTGAGGTGAGGAGTTCGAGACCACCTGGCCAACATGGCAAAACCTATCT  
 CTACTAAAAACACAAACATGAGCCGGGTGTGTGGTGGGAGCCTGTAATCCAGCTACTC  
 GGGAGTCTGAGACAAGAGAATCACTTCAACCCGGGAGGGCGAGGTGTCAGTGAGCCAAGA  
 TCGCGCCATTGCACTCCAGCTGGGCAACGAGAGCGAAACTCCGTCTCAAAAAAAAAAA  
 AAAAAAAAAAGATTACTTTCTTTTATCATTCCTTTATCTTTAAAGCTTTCTTGCACTCA

14629 TGTTTATCTCCAAATGAATGCAGAAATACTAATTATCTTTTTTCTGGTTCTGGGAACA  
 CAGAATTCTAGCGGCTTGTGGAGCCATTTCCCTGGAGCCATGGGGCTCCAGGTCTTT  
 CCTGTGTCTTATTTTTTACGAATTTTTTCTTTTTTGAGACAGGATCTTGCTCTGACTC  
 CCAGGCTGGAGCAATCATCGCTCACTCAAGCGATCTCCACCTCAGGCTCCACGTA  
 GCTGGGACTACAGGTGAGCACCACCATCTGGCTAATGTTTTTAAATTTTTTGTAGGG  
 [G, A]  
 TGGGTCTCACTATGGTGCCAAGACTAGTCTTAAACTCCTGGCCTCAAGAGTTCTCCTG  
 CCTTGGCCTCCAAAGCACTGGGATTACAGGAATGAGCCTCATGCTGGGCTTTGCTGG  
 CGTCTTCAGAGCCCTAGGTACAGGGCCAGCCTGGCGCCCTGCCGCAAGCTTATCTTAAA  
 GCTGGGACCACAACATGCATACCTGCAGCCGGGCCCCGGGGCCAGAGGGCTTTGAGGCAGC  
 ATTTCTCAGCCTTTTAGACACACACTCTGTTAACCCCATCCTGTGTCTCTGATAATCTT

14698 TTCAATTTTTACGAATTTTTTCTTTTTTGAGACAGGATCTTGCTCTGACTCCAGGCTG  
 GAGCACAATCATCGCTCACTCAAGCGATCTCCACCTCAGGCTCCACGTAGCTGGGAC  
 TACAGGTGAGCACCACCATCTGGCTAATGTTTTTAAATTTTTTGTAGGGGTGGGGTC  
 TCACTATGGTGCCAAGACTAGTCTTAAACTCCTGGCCTCAAGAGTTCTCCTGCCTTGGC  
 C  
 [T, A]  
 CCCAAAGCACTGGGATTACAGGAATGAGCCTCCATGCTGGGCTTTGCTGGCGTCTTCAG  
 AGCCCTAGGTACAGGGCCAGCCTGGCGCCCTGCCGCAAGCTTATCTTAAAGCTGGGACC  
 ACAACATGCATACCTGCAGCCGGGCCCGGGGCCAGAGGGCTTTGAGGCAGCATTTCTCAG  
 CCTTTTAGACACACACTCTGTTAACCCCATCCTGTGTCTCTGATAATCTTCTGTGATC  
 C

16095 AATACCTGTCCCCTGCGGTGACCTGGATCTGCTAACCTCCACCCCTGCCTAGACTGTGGA  
 AGGATTGCTGGAAGGGTCTCAGTTGCACAGACCAGGAACTGAGGCCACAGAGGCAGGT  
 GTCCGGTTGTTTGCAACCTCTCAGCCTGTGCTAACCCCAATTGTTGAGAGAGGCCCTGA  
 AACCTCTCCTCTGGGCGCCCCAGGTGACTGCCCCAGCCTCAAGGGCTGCCTCTGTTGC  
 AGGAAAGACGAGTCAAGGCTACTTCCCGTCCATGTACCTGCAAAAGTCAGGGCAAGAC  
 [C, T, G]  
 TGTCCAGGCCCAACGCCAGATCAAGCGGGGGCGCGCCCGCAGGTAAGCGGGGTCC  
 CCGGGCTGGGCGGGTGCAGCGGGGCGCACCAGGGTTCGCTCTGTCTAGGCCATAGCT  
 TGGCAGTCCCGGGCGGGGGCTCTCAGCCTGGCAGGAGAGGCAGGACCTCAGGGGGAA  
 AGGGGCTGGACGCGCTGGCCGCGGTGTGGGGTGGACGGGGGCGGAAGGAAAGCGGCG  
 ATGCCCGGGGGCTTTGGGGATGGGCACTCAGGGGGGCTCCCCGAGAGGGGGACGACAG

16266 GAGCCCTGAAACCTCTCCTCTGGGCGCCCCAGGTGACTGCCCCAGCCTCAAGGGCTGC

FIGURE 3M

CTCTGTTGCAGGAAAGACGACGTACAGGCTACTTCCCGTCCATGTACCTGCAAAAGTCA  
GGGCAAGACGTGTCCAGGCCCAACGCCAGATCAAGCGGGGGCGCCGCCCGCAGGTAA  
GCGGGGGTCCCCGGGGCTGGCGGGGTGAGCGGGGCGACACGGGTTCGCTCTGTCTA  
GGCCATAGCTTGGCAGTGCCGGGGCGGGGGCTCTCAGCCTGGCAGGAGAGGACCCCT  
[C,T]  
ACGGGGGAAAGGGGCTGGACGCGCCTGGCCGCGGTGTGGGGCTGGCACGGGGGCGGAAGG  
AAAGCGCGATGCCCGGGGGCTTTGGGGATGGGCAGTCCAGGGGGGCTCCCCGGAGAGGG  
GGACGACAGACCGAAGGCTGGTGAGGGGCGTGAAAACCGCCAGGCTCTGCTGCAGGGC  
AAGGGTCCTTGTCTGTACGGGGGACCCGCCTTGTGCCCGGGGTCTGTGCACTAC  
CGGCCCCCTACTGCCCCCACTTCTCGGACCAGGGGTGCCCATCTGAGTCCCTGGGGG

16629 AGCGGCGATGCCCGGGGGCTTTGGGGATGGGCAGTCCAGGGGGGCTCCCCGGAGAGGGGG  
ACGACAGACCGAAGGCTGGTGAGGGGCGTGAAAACCGCCAGGCTCTGCTGCAGGGCAA  
GGGTCTTGTCTGTACGGGGGACGCCCTCTTGTCCCGCGGGGTCTGTGCACTACCG  
GCCCCCTACTGCCCCCACTTCTCGGACCAGGGGTGCCCATCTGAGTCCCTGGGGGACG  
GGCGGCCCTCGGGCTTTGACGACGCCCGCTCCCGCTGGGCCAGGTCTCATCCGCAACG  
[C,T]  
GCACAGCATCCACAGCGGTGCGGAAGCGCCTCAGCCAGGACGCCTATCGCCGCAACAG  
CGTCCGTTTTCTGCAGCAGCGACGCCCGCAGGCGCGGGACCGCAGAGCCCCGGGAG  
CCCGCTCGGTGAGTGACGCGGGAGAGGGCAGGAAGGGCAAGCCCTAGGGGCGGAGTCAGC  
GGGAGAGCGGGGCGAGGACAGGGCCAGAGTAGCGGGGCGGGACAGAGGGCGGAATCA  
GAGGGAGAGGCGGGGACTGAGGCGGGGCGAGGAGGAGCCAGCGCTAGGGGCGGAGC

16642 GGGGGCTTTGGGGATGGGCAGTCCAGGGGGGCTCCCCGGAGAGGGGGACGACAGACCGAA  
GGCTGGTGAGGGGCGTGAAAACCGCCAGGCTCTGCTGCAGGGCAAGGGTCTTGTCTGT  
GACGGGGGACGCCCTCTTGTCCCGCGGGGTCTGTGCACTACCGGCCCTACTGCC  
CCCCCTTCTCGGACCAGGGGTGCCCATCTGAGTCCCTGGGGGACGGGGCGCCCTCGGG  
CTTTGACGACGCCCGCTCCCGCTGGGCCAGGTCTCATCCGCAACGCGCACAGCATCCA  
[C,T]  
CAGCGGTGCGGAAGCGCCTCAGCCAGGACGCCTATCGCCGCAACAGCGTCCGTTTTCTG  
CAGCAGCGACGCCCGCAGGCGCGGGACCGCAGAGCCCCGGGAGCCGCTCGGTGAG  
TGCAGCGGGAGAGGGCAGGAAGGGCAAGCCCTAGGGGCGGAGTCAGCGGGAGAGGGGG  
CCAGAGGCGGGGCGAGTAGCGGGGCGGGACAGAGGGCGGAATCAGAGGGAGAGGCGG  
GGACTGAGGCGGGGCGAGGAGGAGCCAGCGCTAGGGGCGGAGCGCATCCCTAAGAGG

18537 AAAAAAGTAACCTTAGGTGCAGGTGTCTCTGTATTCTACTGAGACCGTGCCCCGGTTAT  
GAGGTTGTACCAGAAAGCAAGTATTCACTATGCACACTATTCACCGCTCACCTAGCATT  
GAAGCCAGCCTGTAGCCTGAAAGCCTTTGCTTTGAGGGCAGGTCTTTCCCCAAATGCAG  
ACACGAAGGTGCAAGTGAAGCTGCCAGTCTTGCAAAAGATGTAACCTGTACGAAGGCC  
ACGAGTGGCAGGGAGAGCTGTCCACATTTGCGGAAGTGGCTATGTGAGGACGGGGGAGG  
[C,T]  
GGGTCCCTTAGAGATAAGAGACAATCATAAGGGGAGATATCAGAGAAAATCGTAAGGGGA  
GCAGATGGTTGTCAAGAGAATAGGCTGACCATCGAAGGACTGGCAGAAGCTTTCAGAAAA  
CCACTGGACGGCTGGGCACAGTGGCTTAGGCCTGTAATCCAGCACTTTGGGAGGCTGAC  
CAGGTGAATCACTTGAGGTGAGGATTCAGACCAGCCTGGCCAACATGGTGAAACCC  
ATCTCTACAGAAAATATAAAAATTAGCCAGGCGTGGTGGCACAAGCCTAGAATCCCAGCT

18589 CCGGTTATGAGGTTGTACCAGAAAGCAAGTATTCACTATGCACACTATTACCGCTCACC  
CTAGCATTTGAAGCCAGCCTGTAGCCTGAAAGCCTTTGCTTTGAGGGCAGGTCTTTCCCCA  
AAATGCAGACACGAAGGTGCAAGTGAAGCTGCCAGTCTTGCAAAAGATGTAACCTGTCA  
CGAAGGCCACGAGTGGCAGGAGAGCTGTCCACATTTGCGGAAGTGGCTATGTGAGGAC  
GGGGGAGGCGGGTCCCTTAGAGATAAGAGACAATCATAAGGGGAGATATCAGAGAAAATC  
[G,A]  
TAAGGGGAGCAGATGGTTGTCAAGAGAATAGGCTGACCATCGAAGGACTGGCAGAAGCTT  
TCAGAAAACCACTGGACGGCTGGGCACAGTGGCTTAGGCCTGTAATCCAGCACTTTGGG  
AGGCTGACGCGAGTGAATCACTTGAGGTGAGGATTCAGACCAGCCTGGCCAACATGGT  
GAAACCCATCTCTACAGAAAATATAAAAATTAGCCAGGCGTGGTGGCACAAGCCTAGAA  
TCCAGCTACTTTGGGAGGCTGAGG

18720 CGAAGGTGCAAGTGAAGCTGCCAGTCTTGCAAAAGATGTAACCTGTGACGAAGGCCACG  
AGTGGCAGGGAGAGCTGTCCACATTTGCGGAAGTGGCTATGTGAGGACGGGGGAGGGCG  
GTCCCTTAGAGATAAGAGACAATCATAAGGGGAGATATCAGAGAAAATCGTAAGGGGAGC  
AGATGGTTGTCAAGAGAATAGGCTGACCATCGAAGGACTGGCAGAAGCTTTACAGAAAAC  
ACTGGACGGCTGGGCACAGTGGCTTAGGCCTGTAATCCAGCACTTTGGGAGGCTGACGC  
[G,A]

FIGURE 3N

GGTGAATCACTTGAGGTCAGGAGTTCCAGACCAGCCTGGCCAACATGGTGAAACCCCATC  
TCTACAGAAAATATAAAAATTAGCCAGGCGTGGTGGCACAAGCCTAGAATCCCAGCTACT  
TGGGAGGCTGAGG

18782 TGGCAGGGAGAGCTGTCCACATTTGCCGAAGTGGCTATGTGAGGACGGGGAGGCGGGT  
CCCTTAGAGATAAGAGACAATCATAAGGGGAGATATCAGAGAAAATCGTAAGGGGAGCAG  
ATGGTTGTCAAGAGAATAGGCTGACCATCGAAGGACTGGCAGAAGCTTTAGAAAACAC  
TGGACGGCTGGGCACAGTGGCTTAGGCCTGTAATCCCAGCACTTTGGGAGGCTGACGCA  
GTGAATCACTTGAGGTCAGGAGTTCCAGACCAGCCTGGCCAACATGGTGAAACCCCATCT  
[C,T]  
TACAGAAAATATAAAAATTAGCCAGGCGTGGTGGCACAAGCCTAGAATCCCAGCTACTTG  
GGAGGCTGAGG

18841 TCCCTTAGAGATAAGAGACAATCATAAGGGGAGATATCAGAGAAAATCGTAAGGGGAGCA  
GATGGTTGTCAAGAGAATAGGCTGACCATCGAAGGACTGGCAGAAGCTTTAGAAAACCA  
CTGGACGGCTGGGCACAGTGGCTTAGGCCTGTAATCCCAGCACTTTGGGAGGCTGACGCA  
GGTGAATCACTTGAGGTCAGGAGTTCCAGACCAGCCTGGCCAACATGGTGAAACCCCATC  
TCTACAGAAAATATAAAAATTAGCCAGGCGTGGTGGCACAAGCCTAGAATCCCAGCTACT  
[C,T]  
GGGAGGCTGAGG

Chromosome mapping  
Chromosome 19

FIGURE 30